Collective decisions with incomplete individual opinions

Terzopoulou, Z.

Publication date
2021

Document Version
Final published version

Link to publication

Citation for published version (APA):
Terzopoulou, Z. (2021). *Collective decisions with incomplete individual opinions*. [Thesis, fully internal, Universiteit van Amsterdam]. Institute for Logic, Language and Computation.

Collective Decisions
with Incomplete Individual Opinions

Zoi Terzopoulou

Collective Decisions
with Incomplete Individual Opinions
Zoi Terzopoulou

Collective Decisions
with
Incomplete Individual Opinions

Zoi Terzopoulou

Collective Decisions
with
Incomplete Individual Opinions

ILLC Dissertation Series DS-2021-06

INSTITUTE FOR LOGIC, LANGUAGE AND COMPUTATION

For further information about ILLC-publications, please contact

Institute for Logic, Language and Computation
Universiteit van Amsterdam
Science Park 107
1098 XG Amsterdam
phone: +31-20-525 6051
e-mail: illc@uva.nl
homepage: http://www.illc.uva.nl/

These investigations were supported by the Dutch Research Council (NWO) as part of the *Customisable Collective Choice* project funded under the Physical Sciences TOP Grant scheme (grant number 612.001.652).

Cover art: James Grayot & Zoi Terzopoulou. Cover design: Moriz Oberberger
Printed and bound by Online-Druck GmbH & Co. KG.

ISBN: 978–90–903–4572–7

Collective Decisions with Incomplete Individual Opinions

Academisch Proefschrift

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. ir. K.I.J. Maex
ten overstaan van een door het College voor
Promoties ingestelde commissie
in het openbaar te verdedigen in de Aula der Universiteit
op woensdag 2 juni 2021, te 14.00 uur

door

Zoi Terzopoulou

geboren te Marousi

Promotiecommisie

Promotor:	prof. dr. U. Endriss	Universiteit van Amsterdam
Co-promotor:	prof. dr. K.R. Apt	Universiteit van Amsterdam
Overige leden:	prof. dr. F. Dietrich	Paris School of Economics
	prof. dr. E. Elkind	Oxford University
	dr. D. Grossi	Universiteit van Amsterdam
	prof. dr. H.J.M. Peters	Maastricht Universiteit
	prof. dr. G. Schäfer	Universiteit van Amsterdam
	prof. dr. S.J.L. Smets	Universiteit van Amsterdam

Faculteit der Natuurwetenschappen, Wiskunde en Informatica

To my grandmother Ntina,
who knew how to make the best collective decisions,
usually by only asking herself

Στη γιαγιά τη Ντίνα,
που ήξερε πώς να παίρνει τις καλύτερες συλλογικές αποφάσεις,
συνήθως ρωτώντας μόνο τον εαυτό της

Contents

Acknowledgments

Although I had not imagined that it would be on the topic of collective decision making, the idea of writing a book has been incubating in my mind since I was a kid. So many people have contributed to bringing this dream to reality that I will naturally (yet unfortunately) have to make a subjective selection about whom to mention here.

First, and above all others in my list, is Ulle Endriss. Ulle: You have taught me everything I know about what it means to be an academic—a respectful, supportive, funny, organised, active, conscious, creative, intelligent, social (conditionally), and opinionated one. Thank you for always trusting me with definitions and proofs, money and travels, ideas and ambitions, stress and planning, during plenty of difficult (personally and globally) life events. Thank you for guiding me, but also for listening and thinking along, no matter what questions I would find myself drawn to. The past six years in Amsterdam (almost five of which I spent under your supervision) have shaped me irreversibly, and each day I find myself a step closer to being the person I want to be—I owe a considerable part of this to you. I have to accept that I will not be able to convey the full extent of my gratitude within a few lines of text, so I will move on.

I am deeply honoured that researchers whom I have been admiring for a long time and who have profoundly inspired my work have agreed to be in my committee. Krzysztof Apt, Franz Dietrich, Edith Elkind, Davide Grossi, Hans Peters, Guido Schäfer, and Sonja Smets: I sincerely thank each and every one of you.

Much of my favourite material would be missing from this thesis if it weren't for two wonderful collaborators, namely Justin Kruger and Panagiotis Terzopoulos, the latter of which also happens to be an amazing brother. Justin: Thank you for the afternoons playing on the whiteboard with a view of the Eiffel tower, and for transmitting pure research enthusiasm. Panagioti: Thank you for joining me in this fascinating adventure, and for opening up a whole new, programming world to me. To all my coauthors whose work does not appear in this thesis: Thank you for everything—it was fun!

Special thanks to Stéphane Airiau, Franz Dietrich, Jérôme Lang, and Nicolas Maudet, for facilitating my stay in Paris during two precious months in spring 2019, and to Svetlana Obraztsova and Marija Slavkovik for hosting two highly enjoyable and productive research visits in Singapore and Bergen, respectively. I feel fortunate to be calling myself a member of the computational social choice community more broadly—everyone has been genuinely welcoming from the very beginning, providing an ideal work environment from which I hope to profit for many years to come.

To the members of the COMSOC group in Amsterdam (Adrian, Arianna, Arthur, Julian, Ronald, Sirin, Simon): You are the reason group meetings were warmly anticipated, even recently when they could only take place behind screen barriers. You have individually, and collectively, played a role in some of my most vivid and important ILLC memories (tasting cake on gloomy Wednesdays included). Thank you for the jokes, the confidence boost, the gossip, and last but not least, the serious scientific input. Thank you also for assisting during the development of many of the examples and pictures in this thesis—the pages that follow would be less pretty without you.

Several more people that I had the chance to meet in the ILLC will always occupy a special place in my heart, two of which also served as my flatmates. Levin: Thanks for your smiles and for always being available to share a worry. Ana: You radiate optimism, and have been a comforting presence when I needed it the most; thank you! Two talented designers and good friends should also be acknowledged here—not only for contributing to the unique Sunday atmosphere in our flat, with brunches and pizza, but also for sprinkling my work with colour, in different occasions. Moriz: Thank you for helping me create a beautiful cover for this thesis. Karin: Thank you for being eager to experiment with designing a poster for abstract research on judgment aggregation, and for offering me the opportunity to give my first podcast interview.

An extremely difficult task was also assigned to my paranymphs, who had to plan a PhD graduation ceremony under the tremendous amount of uncertainty that the world is experiencing at the moment. Laura: Thank you for being there when I took the wise decision to embark on this journey of social choice, and for staying (t)here despite the distance and our busy schedules. Sirin: Thank you for providing some distraction in the office, reminding me that there are more things in life besides calculations of voting outcomes, for introducing me to new hobbies during the lockdown, for brainstorming a lot (maybe more than necessary), and for letting me vent relentlessly; I would reorganise a seminar with you, and this, I believe, is the greatest compliment.

To my friends around the globe, and to my loving family in Greece: Thank you for giving me good reasons to forget about the contents of this thesis—spending my breaks in your company was essential to keep myself motivated and happy.

James: You have witnessed all dimensions of my growing during the PhD, and have always known what could make my experience a bit more pleasant. I am grateful that you are on my side, in the moments that make sense, and in those that do not. Your life perspectives and problem-solving approaches are invaluable—thank you!

Μαμά, μπαμπά, και Παναγιώτη: Σας ευχαριστώ για την αέναη εμπιστοσύνη στις ικανότητές μου, για την υποστήριξη στις επιθυμίες μου, και για τη θάλασσα που πάντα είστε έτοιμοι να προσφέρετε. Δε θα ήμουν εδώ (αλλά ούτε και πουθενά αλλού) χωρίς εσάς. Είστε η φωλιά που είχα ανάγκη για να πετάξω.

Chapter 1

Introduction

The defense date of this thesis has been arranged for June 2nd, 2021. As is customary, the committee members were asked to express their preferences over several alternative dates in June, and a consensus was found amongst them. This example is far from being an isolated one. Myriad other processes of collective decision making take place in our everyday life. From finding the best compromise when choosing a restaurant to have dinner with friends, to selecting the US president, collective decisions come in all shapes and sizes: they can materialise in a small or large scale, and they can be attached to low or high stakes.

Considering the prevalence of collective decisions in all functioning societies, it is not hard to believe that their systematic study has been ongoing for a long time—characteristically, it was taken upon by a marquess during the French Revolution: Nicolas de Condorcet. Condorcet and his antagonists (including, notoriously, another French mathematician of that period, Jean-Charles de Borda) found interest in disputing the methods that the French Academy of Sciences used to elect its members, planting seeds for the field that is now known as *social choice theory* (Arrow et al., 2002).

The second wave of social choice theory was instigated by the famous work of Kenneth Arrow around two centuries later, in the 1950's, with other representatives of that era being Amartya Sen and Duncan Black. Arrow came up with a simple, yet tragic answer to the disagreements of Condorcet and Borda, who were trying to single out the best voting method for making collective decisions: Not only did he prove that none of their suggested methods fit the role, but he also showed that the only method that did has to take the form of a dictatorship of one specific individual. Gradually, the field became saturated with ample results of that kind, painting a pessimistic picture for the future of social choice.

Luckily for this thesis, we are currently experiencing a renaissance in the area. Building upon social choice theorists that traditionally were trained in the disciplines of economics, mathematics, and political science, the growth of computer science and artificial intelligence has opened the way for many more research expeditions on collective decision making, giving birth to the sub-field of *computational social* *choice* (Brandt et al., 2016). Today, models of social choice must tackle a twofold, unprecedented challenge: First, accounting for decisions made by artificial agents in addition to human ones; second, providing sufficient flexibility for decision processes that take place in online forums rather than in physical environments. But besides bringing to the surface further problems demanding a solution, computer science has also shed new light on several classical topics of social choice theory. For instance, Gibbard (1973) and Satterthwaite (1975) independently showed that the participants of many collective decision tasks will have an incentive to lie about their personal views in order to achieve a preferable outcome. Yet, it is now proven that such behaviour may be computationally expensive, and thus arguably difficult to achieve in practice (Faliszewski et al., 2010).

Coming back to the matter of choosing a defense date for this thesis, let me disclose that expressing my own preferences over the possible options was not an easy affair. I found dates at the beginning of June attractive, in that they were promising of a long summer break; but I also deemed the dates in late June desirable, since they would afford me a more comfortable writing period in spring. If I had to strictly rank June 5th and June 25th, I would not be able to do it. Many of the popular models in social choice theory are not able to account for such situations of *incomplete preferences*, a problem that also is not addressed by most of the modern frameworks in computational social choice. Related considerations constitute the driving force of this thesis, posing a collection of fundamental research questions.

> Incompleteness arises naturally in many scenarios where the members of a group—that need to make a decision together—are not able to form their personal opinions about all issues under consideration. In order for models of collective decision making to be practically appealing in such contexts, they need to allow for incomplete individual opinions and treat them in an appropriate manner. How can collective decisions be produced then? Which of all possible mechanisms that aggregate individual opinions are normatively compelling? Which of them are effective in discovering the true state of a given situation, in incentivising the agents to be sincere, or in achieving several of the above in parallel?

The departing point of our analysis is that agents who abstain on a given issue to be decided upon do not hold any positive or negative opinion about that issue. Similarly, agents who do not report any comparison between two options do not have any preference over them. In that sense, individual opinions are *intrinsically* incomplete.

Incompleteness may stem from several factors that influence the formation of individual opinions. A basic one has been noted, and widely studied, in the area of *multicriteria decision making*: Every issue at stake in a decision scenario can be seen as a bundle of criteria that are satisfied in various degrees. For example, a certain car that an agent may consider buying is associated with a given gas consumption, a given colour, a given seat comfort, and a given price. When contemplating an option, the agent evaluates the criteria, in comparison to other potential options. Informally, we can say that an agent prefers option a to option b if a is better than b on all criteria, that the agent is indifferent between the two options if they are equally good on all criteria, and that the agent considers the options incomparable if a is better on some criteria but b is better on some others.

A different cause for incomplete individual opinions is that the agents may only know the *names* of the options involved in the decision making process, without being aware of all their *attributes*. The severity of this cause is exacerbated nowadays, where multiple decisions are made on the Internet—the amount of available options in online settings exceeds the ability of any single agent to critically evaluate them. In this thesis, we wish to take the very fact that some of the individual opinions may be incomplete into account for the produced collective decisions. We consider incompleteness as an elementary feature of an agent's opinion, including cases where it could potentially be alleviated in some decision making context (e.g., two board games that I have never played may be part of a very clear ordering for you), and cases where the agent's opinion is theoretically uncompletable (e.g., because it tries to compare apples with oranges),

We will see many more examples of collective decisions, often relatable ones, but sometimes abstract. We will follow a group of three friends, Ann, Bea, and Cal, while they are making decisions together in all kinds of situations: They will be reviewing scientific papers, examining patients, visiting hotels, and being members of juries. We will also eavesdrop on people deciding whether fresh fruit should be offered on a college campus, after forming agreements with farms of dubious moral quality.

All collective decisions that we will explore will be formalised within one of two mathematical frameworks: *judgment aggregation* and *preference aggregation*. The former encodes decisions on binary issues that are possibly connected through logic (such as the decision to organise an online event, which requires a decision about a web conferencing app); the latter captures decisions about a set of alternatives on which the members of a group have preferences (for example, the selection of movies to include in a family's watch-list). Let us elaborate on these two frameworks a bit further.

Judgment Aggregation. The need for a general and flexible model of collective decisions led to the introduction of judgment aggregation less than twenty years ago (List and Pettit, 2002)—a framework that has its origins in the work of Wilson (1975) on aggregating binary values. In judgment aggregation, agents express *yes*/*no* opinions on interrelated issues, which may range from topics of a referendum and evidence of a juridical case to the assessment of the room temperature. As such, this framework transcends disciplinary boundaries, providing a rich playground for researchers in political science and law, as well as in multiagent systems and artificial intelligence.

An infamous way to introduce judgment aggregation is through the *discursive dilemma* (Pettit, 2001)—we will now explain it in yet another variation. Imagine three flatmates having to reach a collective agreement on three issues: whether the flat's expenses of last month were reasonable, whether watching TV shows is a fun hobby, and whether a subscription for Netflix should be bought. They all agree that the best way to deal with the situation is to buy the subscription *if and only if* the first two conditions are satisfied. One of the friends finds the expenses reasonable but dislikes TV and thus is against Netlfix. Another one of the friends loves watching TV but thinks that they already spent enough money on the flat last month, so she also rejects the idea of Netflix. The third friend agrees with the first on that the expenses were quite low, and with the second on that TV is great, and thus supports a Netflix subscription. Overall, for each of the two conditions that were set by the flatmates, a majority is in favour. Yet, a majority is also opposed to the subscription. Although the friends individually followed the consistency requirements of the decision making problem, a collective outcome based on majority did not. Achieving consistent collective decisions is a common struggle of scholarly work in judgment aggregation, which usually assumes that all agents hold a concrete opinion about all issues in question.

But the theme of including incompleteness in studies of judgment aggregation is not new. Prior work in the area has already considered incomplete judgments, both at the collective and at the individual level. In the former direction, the assumption that an aggregation mechanism has to produce a complete collective decision has been relaxed by Gärdenfors (2006), Dietrich and List (2008), and Dokow and Holzman (2010) in the hope of circumventing typical impossibility results of the field stating that there does not exist any reasonable aggregation mechanism satisfying simultaneously a handful of desirable properties. Allowing agents to possibly abstain on some of the issues has been explored by the same authors as well, again in the light of some positive news regarding possibility results (with no great success). Additional papers have specifically focused on the design of new aggregation mechanisms tailored to incomplete inputs. For instance, Slavkovik and Jamroga (2011) have treated an abstention as a vote bearing a third value (in parallel to positive and negative values) and have defined a class of distance-based rules in this setting. Jiang et al. (2018) have constructed an aggregation mechanism that hinges on a hierarchy over the agents and have shown that this mechanism satisfies a number of desirable properties in scenarios with incompleteness.

Preference Aggregation. There is no doubt that general frameworks like judgment aggregation are valuable, since they highlight common properties amongst the many situations they capture (in our case, situations of collective decision making). But narrower models come with their own benefits too; they provide structure that is important for technical observations, and they also allow for the more careful inspection of properties that may be specific to given contexts. Preference aggregation is such a model, narrower mathematically speaking, yet an interesting and widely established one. Settings of preference aggregation can be simulated within judgment aggregation: A *yes* judgment on a pair of alternatives a and b translates to a preference of a over b, while a *no* judgment is interpreted as a preference of b over a.

A collective decision based on individual preferences over a set of alternatives may have various purposes, amongst which is selecting the single favourite alternative of the group on the one hand (which commonly takes place in *voting*), and yielding a collective ranking over all alternatives on the other hand. The aggregation functions that can produce collective outcomes and that are employed in the economics literature are called *social choice functions* in the former case, and *social welfare functions* in the latter. In this thesis, we will simply use the terms voting and preference aggregation to refer to collective decision making that involves individual preferences instead of judgments. The type of aggregation mechanisms used will be clear within our results.

Preference aggregation does not escape precarious instances of collective decision making as the discursive dilemma that we demonstrated in judgment aggregation—the much older counterpart of the discursive dilemma in preference aggregation is due to de Condorcet (1785), and is known as a *Condorcet cycle*. Consider three friends, with preferences over three alternatives a, b, c. One prefers a to b, b to c, and a to c, another one ranks b first, c second, and a third, while the last one has c as her favourite option, followed by a and b. What is the preference of the majority then? It is easy to see that two out of the three friends prefer a to b, while another pair prefers b to c, and a third pair prefers c to a, implying a cyclic ranking.

A plethora of articles explore the circumvention of Condorcet cycles, typically by inspecting their frequency through simulations (e.g., Gehrlein, 2002), or by imposing constraints on the possible preferences of the agents. The work that is closest to this thesis as far as assumptions are concerned emanates from economics: For incomplete (partial) individual preferences, Fishburn (1970) has studied conditions under which simple majority leads to transitive collective preferences; his article is a nice starting point for the reader who is interested in discussions that were taking place in the social choice circles during the period following Arrow's theorem.

Some authors in computational social choice have also explored the aggregation of incomplete preferences, but they have commonly handled incompleteness as an unavoidable inconvenience; either they try to escape from it by considering compatible ways of completing the relevant preferences, or they analyse in depth the ways that negative results of the field are extended under its presence (see, e.g., Konczak and Lang, 2005; Pini et al., 2007, 2008, 2011; Xia and Conitzer, 2011). As we already explained, here we adopt a radically different, constructive attitude towards incompleteness (an attitude that is shared, for example, by Zwicker (2018), who has proposed multiple generalisations of the voting rule known as Kemeny for special domains of preferences).

Throughout the thesis, some text appears framed by a vertical line on the left:

> Here, we comment upon technical or conceptual details that are not necessary for the general understanding of our ideas and results. The reader is welcome to skip this.

Figure 1.1 provides a graphical overview of the main body of this thesis, with arrows indicating dependency relationships between the chapters (information from the chapter at the starting point of an arrow is used in the chapter at the ending point of the same arrow). A synopsis of these chapters' content comes next.

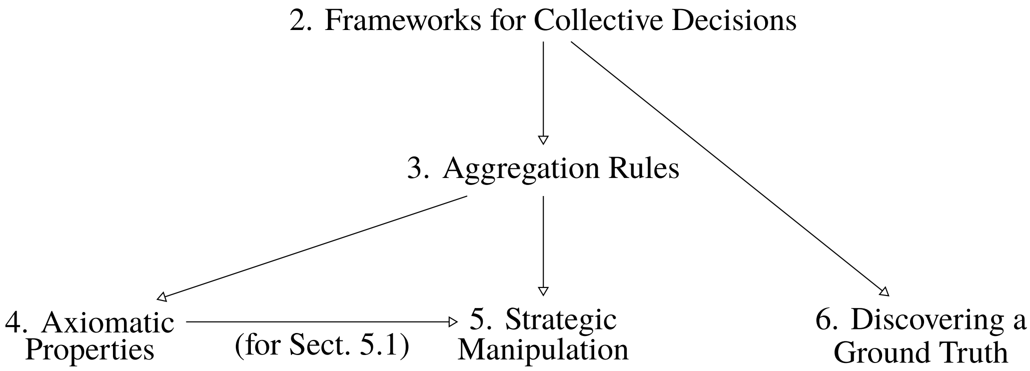


Figure 1.1: Thesis overview and dependency graph.

We promised that this thesis will lay the groundwork for manifold collective decisions that can be made when the members of a group hold incomplete individual opinions. Our endeavour begins with **Chapter 2**, where we provide the essential background for collective decisions in formal contexts. We explain the frameworks of judgment aggregation and of preference aggregation in mathematical terms, and we provide all essential definitions for notions that will frequently appear in the remainder of the thesis (such as preferences and judgments, but also agents, alternatives, and aggregation rules). Chapter 2 elucidates what it means to study collective decision making within (computational) social choice.

1.1 Collective Decisions

Suppose that we have a framework that allows the members of a group to report the incomplete opinions they might possibly have. What kind of mechanisms can be employed for making a collective decision then? For example, how could I, systematically, aggregate the preferences of my committee members in order to decide on a final date for my defense? **Chapter 3** makes several suggestions by means of *aggregation rules* for incomplete preferences and judgments.

We introduce a class of rules that assign weights to agents depending on the size of their reported opinions, a different class of rules that examine whether the number of agents that are in favour of a given option exceeds a fixed threshold, and finally a class of rules according to which alternatives get scores based on their positions in an agent's preference. Chapter 3 (based on the work by Terzopoulou et al., 2018; Terzopoulou and Endriss, 2019a; Kruger and Terzopoulou, 2020; Terzopoulou, 2020; Terzopoulou and Endriss, 2021) thus offers a rich collection of decision making mechanisms for agents with incomplete opinions, all of which are paired with intuitive definitions and natural motivations—however, this chapter does not provide any information about which of the recommended mechanisms are actually *good* to use (in fact, it does not even discuss what it means for an aggregation rule to be a good one).

1.2 Good Collective Decisions

Chapter 4 responds to the question of what a *good* aggregation rule is, for situations of incompleteness. We take a normative stance, and propose a number of properties that an aggregation rule should satisfy in order to be compelling. Such properties are known to the social choice community as *axioms*, and our method as the *axiomatic method*. The novel element of our work consists in incorporating incompleteness; not only in the definitions of the axioms themselves, but also in the axiomatic characterisations of our aggregation rules (we say that an aggregation rule is characterised by a set of axioms if it is the unique rule that satisfies them). For example, suppose that after becoming aware of the preferences of my committee members, I simply choose the favourite date of my supervisor for my defense. The inappropriateness of this mechanism can be easily understood, since only a tiny part of the information provided by the members of the group is used. Even more straightforwardly, this mechanism violates the axiom of anonymity: had the preferences of the committee members been reported to me anonymously, I would not be able to pick the date that my supervisor favoured.

In the axiomatic analysis of Chapter 4, we extend known axioms for the standard case of complete individual opinions (like the axiom of anonymity above), but we also construct new axioms that are customised for contexts of incompleteness. Many aggregation rules presented in Chapter 3 are characterised via our axioms, and are thus justified through them. However, we also show that certain axioms can never hold simultaneously within specific classes of aggregation rules, which stresses that there exist relevant axiomatic properties that are mutually incompatible. The results of this chapter have been included in a number of publications: Terzopoulou et al. (2018), Terzopoulou and Endriss (2019a), Kruger and Terzopoulou (2020), Terzopoulou (2020), and Terzopoulou and Endriss (2021).

1.3 Sincere Collective Decisions

Even if, thanks to Chapter 4, good aggregation rules can be part of our toolbox for making collective decisions with incomplete individual opinions, the members of a group may be happier with an outcome different from what these rules suggest. For example, consider the hypothetical case where June 2nd is the best option for seven out of the eight committee members, but the second favourite option of my supervisor, while June 22nd is the best option for my supervisor and the second best for the rest of the committee. Suppose that my supervisor gets informed about the preferences of all the committee members, by being included in our email correspondence, prior to asserting his own preference. Suppose moreover he knows that I will use a popular aggregation rule to select the defense date, which would result being June 2nd. If my supervisor thinks strategically, instead of revealing his true preferences, he can pretend that June 2nd is his least favourite option in order to steer the collective decision towards

June 22nd. **Chapter 5** investigates the kind of incentives that arise when collective decisions are made with incomplete individual opinions.

The study of strategic behaviour has a long tradition in social choice theory, and is especially pertinent today, when a significant part of decision making takes place online. Internet users can—and do—express falsehoods, using anonymity to avoid accountability. Further more, allowing agents the freedom to report incomplete opinions comes with the risk of them misrepresenting their sincere preferences or judgments in original ways: by withholding opinions they actually have, by declaring opinions they do not have, or by explicitly altering their opinions. We explore these possibilities with respect to the aggregation rules of Chapter 3 and find that incentivising the members of the group to remain sincere is often—although not always—possible. We also examine processes of collective decision making that occur *in iteration*, granting agents the option to change their submitted preferences in rounds, in view of better outcomes they target. Our question in such settings is whether we can identify processes that are guaranteed to terminate relatively fast—we adopt both analytical and experimental tools to answer it. The results of this chapter rely on published work by Kruger and Terzopoulou (2020) and Terzopoulou (2020), and on work in progress in collaboration with Ulle Endriss and Panagiotis Terzopoulos.

1.4 Correct Collective Decisions

Collective decision making is frequently linked to the search of a compromise between the agents in a group, as was the case for the examples we have seen so far. A different, prominent domain where collective decision making is essential concerns the search of a *ground truth* in a given problem: It is meaningless to look for the objectively best date for a PhD defense, but it is imperative to know whether a vaccine objectively works against a virus. Proposing collective choice methods for the latter kind of situations is part of the *epistemic approach* to social choice theory.

Chapter 6 complements Chapters 3, 4, and 5, and examines how groups can make decisions that have high probability to reflect the true answer of a question given that their members are more accurate in finding that answer than a random guesser. Bearing in mind that the source of incompleteness of the agents' opinions may also have an effect on their accuracy, we ask what is the optimal aggregation rule to apply in order to combine the information provided by the agents, as well as what is the ideal way to distribute tasks to the group (in which case a tradeoff emerges between the quality and the quantity of the desired information). The results presented in this chapter are due to Terzopoulou and Endriss (2019b).

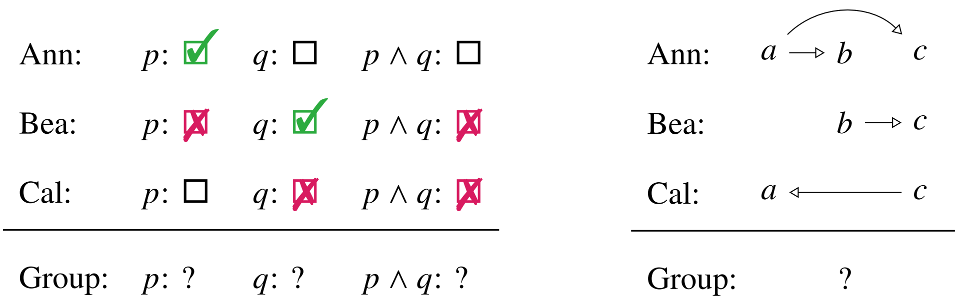
The thesis concludes with **Chapter 7**, where the reader can find a summary of our most important results concerning incomplete opinions in collective decision making, as well as the trailhead to further open questions.

Chapter 2

Frameworks for Collective Decisions

This chapter provides the formal background for this thesis. We review all relevant notation and terminology concerning models of collective decision making. The reader only needs to be familiar with some basic propositional logic.

We already know that different formal frameworks are of use in scenarios where collective decisions are to be made, depending on the kind of opinions the members of a group hold, as well as on the type of collective outcome that is sought. When the opinions are binary *judgments* over logically interconnected propositions (e.g., $p, q, p \wedge q$), then our relevant framework will be *judgment aggregation* (see Section 2.1); when the opinions are *preferences* over a set of alternatives (e.g., a, b, c), we will work within *preference aggregation* (presented in Section 2.2). Figure 2.1 below exemplifies the two frameworks in an abstract manner, for a group consisting of Ann, Bea, and Cal.



(a) Judgment aggregation (for the propositions $p, q, p \wedge q$).

(b) Preference aggregation (for the alternatives a, b, c).

Figure 2.1: Two frameworks for collective decision making.

Starting off, we consider all *potential agents* that may ever participate in a scenario of collective decision making. These, for example, may be conceived as all adults that are simultaneously alive on earth, or as all artificially intelligent bots that are designed by such adults. We define the *superpopulation* $N^\star$ to be a countable set of size $|N^\star| \geqslant 2$ that includes all potential agents. The superpopulation can be assumed to be either finite or countably infinite (in the latter case, the reader may think of it as the set of natural numbers $\mathbb{N}$). The finiteness assumption is realistic if our goal is to build mechanisms for collective decision making that can operate at a specific moment in time—then, $|N^\star|$ simply denotes the number of the existing agents at that moment; on the other hand, assuming an infinite superpopulation is important if we want to have mechanisms that allow for an arbitrarily large number of agents to participate in a collective decision (for instance, as the population of earth increases, we cannot guarantee that fixing a superpopulation of size 7.8 billion will be sufficient tomorrow). We will clarify our assumptions about the cardinality of $N^\star$ whenever these play a role in our proofs.

Each concrete problem of collective decision making involves a finite *group of agents* N, selected from the superpopulation, who hold and report their opinions about some topics in question. For example, in Figure 2.1, the group consisted of three agents: Ann, Bea, and Cal. We will use natural numbers (1,2, etc.) to denote agents, and the letter n for the size of a group. The topics about which a group engages in collective decision making may regard the ranking of certain alternatives (like in voting and preference aggregation) or the acceptance/rejection of given propositions (as is the case in judgment aggregation).

Our key aim is to propose appropriate methods for making collective decisions, given the individual, and frequently diverting, opinions of the agents in a group. To that end, the main protagonists of our work are various *aggregation rules*. An aggregation rule F is a function, that takes as input any collection of individual opinions and outputs a set of collective opinions. For example, an aggregation rule F should yield a collective decision given the opinions of Ann, Bea, and Cal that we saw in Figure 2.1.

$$F: \quad \text{(Ann's opinion, Bea's opinion, Cal's opinion)} \quad \mapsto \quad \text{collective decision}$$

Depending on the kind of opinions present in a context of collective decision making, an aggregation rule may be referred to as a "preference aggregation rule" or a "judgment aggregation rule"—usually, we will simply call it a "rule".

2.1 Judgment Aggregation

In this section, we introduce an elementary model of judgment aggregation that allows for incomplete individual judgments. Our framework builds on the existing literature regarding the aggregation of complete judgments (List and Pettit, 2002; List, 2012; Grossi and Pigozzi, 2014; Endriss, 2016).

We assume that all possible (binary, and logically interconnected) issues about which a group of agents may ever have to decide belong to the *superagenda* $\Phi^\star$, which is a countable set of formulas in propositional logic. The superagenda captures the pool of all issues with which collective decisions will ever be concerned. For instance, it can simply be the (countably infinite) set of all formulas of propositional logic.[1] In a concrete scenario of judgment aggregation, the agents judge issues drawn from a finite and non-empty subset of the superagenda, called the *agenda*. The agenda Φ contains two propositions, φ and $\neg\varphi$, for every issue $\widetilde{\varphi}$ at stake. Slightly abusing notation, we will assume that double negations cancel each other: $\neg\neg\varphi = \varphi$. We will routinely use the letters p, q, etc. for atomic propositions.

Given an agenda $\Phi \subseteq \Phi^\star$, each agent i in the group is assumed to hold a judgment set (or simply judgment) $J_i \subseteq \Phi$, which formally captures her opinion on the issues under consideration (that is, $\varphi \in J_i$ denotes that agent i has a positive—*yes*—judgment on $\widetilde{\varphi}$; similarly, $\neg\varphi \in J_i$ means that she has a negative—*no*—judgment on $\widetilde{\varphi}$). For convenience, we will often depict a *yes* (*no*) judgment on an issue $\widetilde{\varphi}$ as a *yes* (*no*) on the relevant positive proposition φ. Recall Figure 2.1a: Cal has a positive judgment on issue $\widetilde{p}$ and a negative judgment on $\widetilde{q}$ and $\widetilde{p \wedge q}$, so his judgment set is $\{p, \neg q, \neg p \wedge q\}$.

Individual judgments J are *logically consistent* sets of propositional logic (e.g., an agent cannot be positive on φ and on ψ but negative on $\varphi \wedge \psi$), but not necessarily *complete*: There may exist some issue $\widetilde{\varphi}$ such that $\varphi \notin J_i$ and $\neg\varphi \notin J_i$ (in this case, we say that i *abstains* on $\widetilde{\varphi}$).[2] Also, we say that agent i *supports*/*accepts* proposition φ if $\varphi \in J_i$, and she *rejects* proposition φ if she does not accept it, that is, if $\varphi \notin J_i$.

$\mathcal{J}(\Phi)$ is the set of all logically consistent subsets of the agenda Φ, and $\mathcal{J}(\Phi)^\bullet$ those that are also complete (e.g., Bea's judgment in Figure 2.1a is complete):

$$J \in \mathcal{J}(\Phi)^\bullet \text{ if } (i)\ J \in \mathcal{J}(\Phi) \text{ and } (ii)\ \varphi \in J \text{ or } \neg\varphi \in J, \text{ for all } \varphi \in \Phi$$

Given the individual judgments of all agents, we have a *profile* $\boldsymbol{J}$ of judgments:

$$\boldsymbol{J} = (J_1, \ldots J_n) \in \mathcal{J}(\Phi)^n$$

We also write $(\boldsymbol{J}_{-i}, J'_i)$ for the profile where agent i reports the judgment J'_i, and all other agents report the same judgments as in profile $\boldsymbol{J}$. We denote by $N^{\boldsymbol{J}}_\varphi$ the set of agents who accept φ in the profile $\boldsymbol{J}$:

$$N^{\boldsymbol{J}}_\varphi = \{i \in N \mid \varphi \in J_i\}$$

Analogously, $N^{\boldsymbol{J}}_{\widetilde{\varphi}}$ is the set of agents who do not abstain on the issue $\widetilde{\varphi}$ in the profile $\boldsymbol{J}$ (that is, $N^{\boldsymbol{J}}_{\widetilde{\varphi}} = N^{\boldsymbol{J}}_\varphi \cup N^{\boldsymbol{J}}_{\neg\varphi}$). We write $n^{\boldsymbol{J}}_\varphi = |N^{\boldsymbol{J}}_\varphi|$ and $n^{\boldsymbol{J}}_{\widetilde{\varphi}} = |N^{\boldsymbol{J}}_{\widetilde{\varphi}}|$.

Now, in order to obtain a collective decision given the judgments of the agents, we need an aggregation rule. In the framework of judgment aggregation, a *judgment*

---

[1]As a technical note, for any cardinal numbers $k, \ell \in \mathbb{N} \cup \{\aleph_0\}$, we write $k \leqslant \ell$ whenever: (i) $k, \ell \in \mathbb{N}$ and $k \leqslant \ell$ in the familiar way of comparing natural numbers, (ii) $k \in \mathbb{N}$ and $\ell = \aleph_0$, or (iii) $k = \ell = \aleph_0$. Moreover, we adopt the following convention: $\aleph_0/2 = \aleph_0 + 1 = \aleph_0$. We use $\aleph_0$ to denote the cardinality of every countably infinite set.

[2]Gärdenfors (2006) and the authors of follow-up papers require deductive closure of (incomplete) judgment sets. We do not make this assumption. Our model is designed to only take into consideration the agents' explicitly declared opinions, which might not be closed under logical implication.

aggregation rule F is a function that maps any profile of judgments $\boldsymbol{J} \in \mathcal{J}(\Phi)^n$ for any agenda Φ and any group N, to a set of collective judgment sets for the same agenda, i.e., to a subset of 2^Φ. Thus, there may be a tie between several "best" judgment sets and these judgments need not be consistent or complete in general. A rule is called *resolute* if it always returns a single outcome—i.e., if $|F(\boldsymbol{J})| = 1$ for all profiles $\boldsymbol{J}$. Whenever a rule F is clearly resolute, we will simply treat $F(\boldsymbol{J})$ as a judgment set, rather than as a set of judgment sets. We will shortly present some examples of judgment aggregation rules, focusing on the special case where the judgments of all agents in the group are complete. For more general judgment aggregation rules that allow for incompleteness, the reader shall have to wait until Chapter 3.

The Special Case of Complete Judgments. In our framework, a group of agents may all still hold, and report, complete judgments. We can then consult the traditional literature for relevant judgment aggregation rules.

The simplest such rules probably are the *quota rules*, which collectively accept a given proposition if and only if a sufficiently large number of agents include it in their judgment sets. Unfortunately, quota rules often lead to logically inconsistent outcomes—they may even simultaneously accept a proposition φ and its negation $\neg\varphi$. For an example, consider the ubiquitous majority rule (a quota rule that collectively accepts a proposition whenever more than half of the agents also accept it) and see Table 2.1, which illustrates the discursive dilemma (discussed in the introduction) for an agenda with the positive propositions p, q, and $p \wedge q$. In Section 3.2, we delve into quota rules in their full generality, also taking into account incomplete judgments.

	p	q	$p \wedge q$
Agent 1:	☑	☒	☒
Agent 2:	☒	☑	☒
Agent 3:	☑	☑	☑
Majority:	☑	☑	☒

Table 2.1: The discursive dilemma for complete individual judgments.

On the other hand, there exist rules that circumvent the challenge of inconsistent outcomes directly by construction. A prominent rule of this kind is the *Kemeny rule*, defined as follows for all profiles of complete judgments in $\mathcal{J}(\Phi)^\bullet$:

$$\mathit{Kemeny}(\boldsymbol{J}) = \operatorname*{argmax}_{J \in \mathcal{J}(\Phi)^\bullet} \sum_{i \in N} |J \cap J_i|$$

The Kemeny rule is known under a number of other names—notably: *median rule* (Nehring et al., 2014), *distance-based rule* (Pigozzi, 2006), and *prototype rule* (Miller and Osherson, 2009). We will discuss several ways of generalising the Kemeny rule to domains of incomplete judgments in Section 3.1.1. For the profile of Table 2.1, the Kemeny rule outputs the set of all three judgments submitted by the agents, tied.

2.2 Voting and Preference Aggregation

In this section, we set the background for collective decisions that involve preferences. The agents compare several alternatives with which they are presented, and express preferences that are based on these comparisons. We include all *potential alternatives* on which agents may ever form preferences in a set $A^\star$, which we assume to be countably infinite. For each specific problem of collective decision making that involves preferences we have a finite set of alternatives $A \subseteq A^\star$, with $m = |A| \geqslant 3$, which we naturally call the *set of alternatives*. We will denote concrete alternatives with the letters a, b, c, etc., and use the letters x, y, z, etc. as variables that range over $A^\star$.

To link this setting with the framework of judgment aggregation that we saw in Section 2.1, consider two alternatives $a, b \in A$. An agent may rank a above b (saying *yes* to the proposition expressing that a is preferable to b), she may rank b above a (corresponding to a *no* to the proposition expressing that a is preferable to b), or she may not compare a and b at all (abstaining on a possible ranking between a and b). Intuitively, preferences can thus be encoded as judgments—for further details on the encoding of preference aggregation within judgment aggregation, we refer to the works of List and Pettit (2004) and of Dietrich and List (2007a). Here we follow a more straightforward route, directly building a framework for preference aggregation.

So, every agent performs pairwise comparisons over the alternatives in A and forms her truthful preference $R \subseteq A \times A$: a binary relation over A. A preference R may be *incomplete*, that is, it may offer no comparison between two alternatives $a, b \in A$ (all preferences in Figure 2.1b are incomplete). Different assumptions can be made regarding other properties of these preferences, giving rise to various preference domains that are worth exploring. In this thesis, we are mostly interested in two such properties: whether the preferences allow the agents to be indifferent between two given alternatives (in which case we will also say that the two alternatives are indistinguishable for the relevant preference), and whether they are transitive (meaning whether xRy and yRz implies xRz for all alternatives $x, y, z \in A$). We will always assume that the individual preferences are *acyclic*: for all alternatives $x_1, \dots, x_k$, if $x_1 R x_2, x_2 R x_3, \dots, x_{k-1} R x_k$, then it does not hold that $x_k R x_1$, unless all alternatives $x_1, \dots, x_k$ are indistinguishable according to R. We thus define four concrete preference domains (see also Figure 2.2):

- $\mathcal{D}(A)$: The set of all all acyclic preferences over A that allow for indifference. We write $R \in \mathcal{D}(A)$.

- $\mathcal{P}(A)$: The set of all acyclic preferences over A that are strict (i.e, do not allow for indifference). We write $P \in \mathcal{P}(A)$.

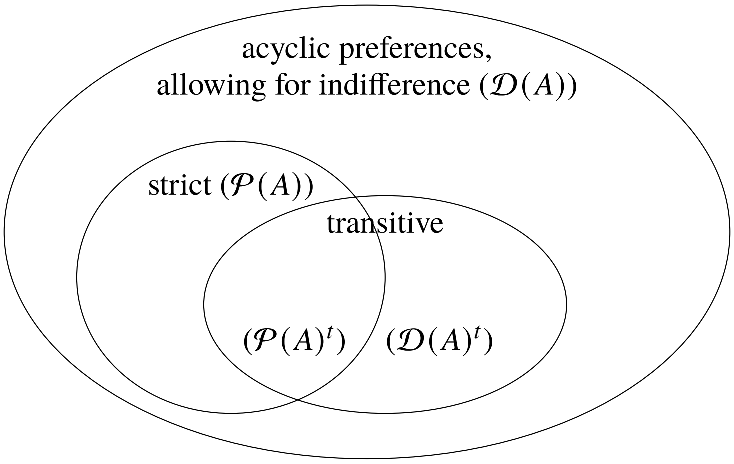


Figure 2.2: The space of incomplete and acyclic preference domains.

- $\mathcal{D}(A)^t$: The set of all acyclic preferences over A that allow for indifference and also are transitive. We write $\succsim\ \in \mathcal{D}(A)^t$.
- $\mathcal{P}(A)^t$: The set of all strict and acyclic preferences over A that also are transitive. We write $\rhd\ \in \mathcal{P}(A)^t$.

Formally, given a preference $\succsim\ \in \mathcal{D}(A)^t$ and two alternatives $a, b \in A$, we write $a{\sim}b$ when $a \succsim b$ and $b \succsim a$, and $a{\succ}b$ when $a \succsim b$ and it is not the case that $b \succsim a$. When $a{\sim}b$, then a and b are said to be *indistinguishable*. When $a{\succ}b$ (or $a \succsim b$), then a is said to be *strongly* (or *weakly*) *preferred to* b. When none of $a \succsim b$ and $b \succsim a$ hold, then a and b are said to be *incomparable*.

We can think of a strict preference $P \in \mathcal{P}(A)$ as a set of ordered pairs:

$$P \quad = \quad \{(a, b) \in A{\times}A \mid a \text{ is ranked above } b\}$$

It will often be convenient to think of pairwise preferences as directed acyclic graphs, with nodes being alternatives, and with a directed edge from a to b if and only if a is considered better than b according to the given preference. See, for example, Figure 2.3.

$$\begin{tikzcd} a \arrow[d] \arrow[dd, bend left] \\ b \arrow[d] \\ c \arrow[d] \\ d \end{tikzcd}$$

The preference $\{(a, b), (b, c), (a, c), (c, d)\} \in \mathcal{P}(\{a, b, c, d\})$.

Figure 2.3: A pairwise preference represented as a directed acyclic graph.

We already stated that we will not always impose the transitivity assumption on the studied preferences, but we did not clearly justify our reason for this choice. Example 2.1 serves as an explanation.

2.1. Example. Imagine you are asked by a travel website to rank different hotels according to your preferences. You have only visited three hotels: *Sandy Cabins* (which is by the seaside), *Luxury Towers* (which is in the city), and *Snowy Chalets* (which is in the mountains). You prefer *Sandy Cabins* to *Luxury Towers* when it is summer, and *Luxury Towers* to *Snowy Chalets* when it is winter. But you would only go to the seaside during summer, and you would only go to the mountains in winter, so comparing *Sandy Cabins* and *Snowy Chalets* does not make sense to you. △

We sometimes write ab to abbreviate (a, b). To talk about the transitive closure of a preference set $\{a_1a_2, a_2a_3, \ldots, a_{k-1}a_k\} \in \mathcal{P}(A)$, we may also write $a_1 \rhd a_2 \rhd \cdots \rhd a_k$.

A *profile* of preferences $\boldsymbol{R}$ captures the preferences of all agents in N, where R_i encodes the preference of agent $i \in N$ (analogously, for preferences represented by different symbols, we will denote profiles by the relevant symbol in boldface):

$$\boldsymbol{R} = (R_1, \ldots, R_n) \in \mathcal{D}(A)^n$$

We write $\langle ab \rangle$ as shorthand for the permutation on A that swaps a and b, and we write $R_{\langle ab \rangle}$ for the preference R with every occurrence of a and b switched. We also apply this notation to profiles: $\boldsymbol{R}_{\langle ab \rangle}$ is the profile $\boldsymbol{R}$ with a and b switched. See Figure 2.4 for an example with a strict and transitive preference profile.

$$\rhd_1 = a \rhd b \rhd c \qquad\qquad \rhd_1 = b \rhd a \rhd c$$
$$\rhd_2 = b \rhd c \rhd a \qquad\qquad \rhd_2 = a \rhd c \rhd b$$

Figure 2.4: The right profile is obtained from the left after swapping a and b.

Given a subset of the agents $N' \subseteq N$, a partial profile $\boldsymbol{R}_{-N'}$ denotes the part of $\boldsymbol{R}$ where all agents besides those in N' report their preferences. By $(R, \ldots, R, \boldsymbol{R}_{-N'})$ we denote the profile where agents in $N \backslash N'$ report the same preferences as in $\boldsymbol{R}$, and all agents in N' report the preference R. Two profiles $\boldsymbol{R} = (R_1, \ldots, R_n) \in \mathcal{D}(A)^n$ and $\boldsymbol{R'} = (R_{n+1}, \ldots, R_{n+k}) \in \mathcal{D}(A)^k$ can be combined to form a new profile:

$$(\boldsymbol{R}, \boldsymbol{R'}) = (R_1, \ldots, R_n, R_{n+1}, \ldots, R_{n+k}) \in \mathcal{D}(A)^{n+k}$$

$N^{\boldsymbol{R}}_{ab}$ denotes the set of agents i for which $ab \in R_i$ in $\boldsymbol{R}$, and $n^{\boldsymbol{R}}_{ab} = |N^{\boldsymbol{R}}_{ab}|$. A *preference aggregation rule* F is a function that maps any profile $\boldsymbol{R} \in \mathcal{P}(A)^n$ for any set of alternatives A and group N to a set of preferences on the same alternatives. Thus, the set $F(\boldsymbol{R})$ may contain several, tied, outcomes. Analogously, a *voting rule* produces a set of winning alternatives (a subset of A), and a *multiwinner voting rule* outputs a set of committees (which are sets of alternatives), given the preference profile of the agents. We will soon discuss some concrete rules for the aggregation of complete individual preferences. Chapter 3 is devoted to the design of such aggregation rules for the more general case of incomplete preferences.

The Special Case of Complete Preferences. The typical voting literature (for a review, consult Zwicker, 2016) concerns agents that hold, and report, preferences in the form of linear (acyclic, strict, and complete) orders, which we denote by $L \in \mathcal{L}(A)$. Given such orders submitted by a group, a very natural way to decide the winner is by assigning points to each alternative depending on the position it appears in the order of every agent, and then select those alternatives with the largest total score across all agents. This procedure is followed by *positional scoring rules*. More precisely, a positional scoring rule $F_{\boldsymbol{p}}$ for profiles of linear orders over m alternatives is induced by a positional scoring vector $\boldsymbol{p} = (p_1, p_2, \ldots, p_m)$ with $p_1, \ldots, p_m \geqslant 0$, $p_k \geqslant p_{k+1}$ for all $k \in \{1, \ldots, m\}$, and $p_\ell > p_{\ell+1}$ for some $\ell \in \{1, \ldots, m-1\}$. When agent i ranks alternative a in the j^{th} position of her preference, then she assigns to it score $p(L_i, a) = p_j$. Given a profile $\boldsymbol{L}$ of linear preferences, we have the following definition:

$$F_{\boldsymbol{p}}(\boldsymbol{L}) = \operatorname*{argmax}_{x \in A} \sum_{i \in N} p(L_i, x)$$

The most famous positional scoring rule for voting with complete preferences probably is the *Borda rule*, introduced by Jean-Charles de Borda (1784). The Borda rule is induced by the scoring vector $(m-1, m-2, \ldots, 0)$. It prescribes that 0 points be assigned to the alternative ranked last by the agent, 1 point to the second to last alternative, and so on, until the top alternative be assigned with $m-1$ points. The *k-approval rules* also are popular positional scoring rules, with the following vectors:

$$\boldsymbol{p}^{k-approval} = (\underbrace{1, \ldots, 1}_{k}, \underbrace{0, \ldots, 0}_{m-k}), \text{ for } 1 \leqslant k \leqslant m$$

For $k = 1$, the corresponding approval rule is known as the *plurality rule* (selecting as winner the alternative that most agents place at the top of their ranking), while for $k = m-1$ the corresponding approval rule is called the *antiplurality rule* (according to which the winner should be among the alternatives ranked last by the agents the least number of times). Note that positional scoring rules are irresolute, i.e., they may return tied winners. See Figure 2.5 for an example.

$$\begin{array}{ccccc} 2 & & 1 & & 0 \\ a & \rhd & c & \rhd & b \\ b & \rhd & c & \rhd & a \\ a & \rhd & b & \rhd & c \end{array} \qquad\qquad \begin{array}{ccccc} 1 & & 1 & & 0 \\ a & \rhd & c & \rhd & b \\ b & \rhd & c & \rhd & a \\ a & \rhd & b & \rhd & c \end{array}$$

Figure 2.5: On the left (Borda): a wins with score 4, b has score 3, and c has score 2; on the right (2-approval): a, b, c all have score 2, and are the tied winners.

In Section 3.3, we describe an extension of positional scoring rules, including the Borda rule, and the k-approval rules, to domains of incomplete preferences.

Positional scoring rules can easily be applied to settings where we need to determine a set of winning alternatives—that is, a committee—as well. If a committee of size k

is required, then we can select those k alternatives with the highest total score. For example, the k-Borda rule for multiwinner voting with linear preferences chooses the winners to be the k alternatives with the highest Borda score (Faliszewski et al., 2017).

Positional scoring rules can also be applied to obtain a collective ranking over the alternatives, by placing on top the alternative with the largest total score across all agents, placing second the alternative with the second largest total score, and so on, by employing some tie-breaking function too, whenever more than one alternative have the same score. A different way to perform preference aggregation over a profile $\boldsymbol{L}$ for the purposes of creating a collective ranking is via the *Kemeny rule*, as follows:[3]

$$Kemeny(\boldsymbol{L}) = \underset{L \in \mathcal{L}(A)}{\operatorname{argmax}} \sum_{i \in N} |L \cap L_i|$$

In Section 3.1.2, we will present a class of rules that generalise the Kemeny rule to domains of incomplete preferences.

[3]One could also first obtain a collective preference ranking, and then select the winners from the alternatives appearing at the top of that ranking.

Chapter 3

Aggregation Rules

The main question with which this chapter is concerned can be phrased as follows:

> How can we aggregate the individual opinions of a group of agents, when these opinions are incomplete?

The first, necessary step in the direction of an answer is to construct aggregation rules that function specifically on domains of incomplete opinions. This is what we are going to do in this chapter, moving between the frameworks of voting, preference aggregation, and judgment aggregation. For most of our definitions, assuming a specific model of collective decision making will be crucial: Some aggregation rules will only be meaningful in preference aggregation and not in judgment aggregation, while for some different rules it may be the other way around. We will always clarify the framework with which we are working, to avoid possible misunderstandings.

The rules we are interested in reduce to known ones from the literature in the special case where all the agents hold complete opinions. This fact directly reflects the spirit of our work, which does not envision to give arbitrary definitions of aggregation rules that would be intriguing just for the sake of their novelty; rather, it provides natural extensions, in more general settings, of rules that are already used and accepted at large within social choice theory. Let us present a high-level overview of this chapter.

Section 3.1 defines rules that take advantage of the idea that, given incomplete opinions, the input provided by different agents during an aggregation process will often vary in quantity—there may be agents who hold opinions about many of the topics under consideration, and there may be agents who do not have an opinion about most of these topics. This situation is relevant for both judgment and preference aggregation (Sections 3.1.1 and 3.1.2, respectively). We construct rules that assign weights to agents based on the size of their opinions.

Section 3.2 deals with quota rules, which are reasonable in the context of judgment aggregation (where a proposition is collectively accepted whenever the number of agents that support it is large enough). One could in theory also formulate quota rules for preference aggregation, by looking at how many agents agree with given rankings between pairs of alternatives. However, this process cannot be epected to lead to collective preferences of a desirable form. For example, quota rules for preference aggregation may induce a collective decision in favour of three separate rankings "a is better than b", "c is better than d", and "e is better than f" that is not useful for practical purposes (not determining any winners, nor any order berween the alternatives).

Lastly, we return to the framework of voting in Section 3.3, where we examine positional scoring rules. Positional scoring rules give scores to alternatives depending on their position in the agents' preferences. When we talk about positional scoring rules in this thesis, we will exclusively refer to voting. An interpretation of scoring rules within judgment aggregation, focusing on complete judgments, has been described by Dietrich (2014). Bringing Dietrich's ideas one step further, to scenarios with incompleteness, we see that the weight rules for judgment aggregation discussed in Section 3.1 in fact constitute a compelling kind of scoring rules.

3.1 Weight Rules

In this section, we present a class of aggregation rules that use weights as their main component. These weights are assigned to the agents and depend specifically on the size of the agents' reported opinions, as motivated in Example 3.1.

3.1. Example. Consider a review phase of some scientific conference, and the reviewers that take part in it. Suppose that one reviewer has read and formed an opinion on only three out of the twenty submissions, while another reviewer has evaluated all twenty of them. Should the same weight be assigned to those two reviewers? Should the reviewer who holds an opinion on some selected few papers be given less, or maybe more, weight? In a similar context, suppose that a group of reviewers consisting of Ann, Bea, and Cal hold the opinions demonstrated in the table below, encoding the acceptance of two papers (p_1 and p_2).

	p_1	p_2
Ann:	☑	☐
Bea:	☒	☑
Cal:	☒	☑

Ann has an opinion only on one of the two relevant papers, which disagrees with the opinion of Bea and Cal who have reviewed both of them. What should the conference chair decide about that paper? On the one hand, if the reviewers spent the same, limited time on their tasks, it is safe to assume that Ann invested all this time on paper 1, so her opinion about that paper can be deemed of higher quality. On the other hand, we may want to weigh more the opinions of the reviewers who have gathered further experience by reviewing more papers. △

This section builds a foundational toolbox to address questions like the ones raised by Example 3.1. We will define weight rules both for preference aggregation and for judgment aggregation, starting from the latter framework.

3.1.1 Judgment Aggregation

The class of weight rules that we are ready to introduce in the framework of judgment aggregation can be thought off as a subclass of what has previously been defined by Dietrich (2014) as the class of *scoring rules*.[1] In this thesis, we choose the distinctive term "weight rules" in order to be able to use the same name for the weight rules that we will later define for voting (in Section 3.1.2).[2]

Our motivating example with the reviewers hinted upon the fact that weight rules will need to determine how much to weigh an agent's judgment, depending on how many propositions she evaluates. To that end, we define a *weight function* $w : \mathbb{N} \to \mathbb{R}^+$, which assigns to every size of a judgment set, $|J_i| \in \mathbb{N}$, a positive real number measuring how a proposition $\varphi \in J_i$ performs from the perspective of holding judgment J_i. We write $w_{|J_i|}$ as an alternative for $w(|J_i|)$.

For instance, w_1 designates the weight corresponding to all judgment sets that are singletons. We may also think of w as a vector of infinite length:

$$\boldsymbol{w} = (w_1, w_2, w_3, \ldots)$$

For example, $\boldsymbol{w} = (1, \frac{1}{2}, \frac{1}{3}, \ldots)$ means that we distribute a total weight of 1 equally across all propositions in J_i, while $\boldsymbol{w} = (1, \frac{1}{M}, \frac{1}{M^2}, ...)$, for a sufficiently large number M, weighs smaller judgment sets substantially more than larger ones. We stress that these weights do not bear upon a specific agent, or the specific propositions evaluated by that agent—the only thing that matters for a weight rule is the size of the submitted judgments. In Example 3.1, using the weight vector $\boldsymbol{w} = (1, \frac{1}{2}, \frac{1}{3}, \ldots)$ entails that the only proposition in Ann's judgment set will get weight 1, while each of the two propositions in Bea's (or Cal's) judgment set will be assigned weight $\frac{1}{2}$.

Then, we will encode the weight of a set of formulas J from the perspective of holding judgment set J_i, by counting the total weight of its members. Concretely:

$$w_{|J_i|}(J) \quad = \quad \sum_{\varphi \in J} w_{|J_i|} \cdot \mathbb{1}_{\varphi \in J_i} \quad = \quad w_{|J_i|} \cdot |J \cap J_i|$$

Suppose again that we use the weight vector $\boldsymbol{w} = (1, \frac{1}{2}, \frac{1}{3}, \ldots)$, and consider the judgment set $J = \{\neg p_1, \neg p_2\}$. In Example 3.1, if Bea is agent i, then we have that $w_{|J_i|}(J) = \sum_{\varphi \in J} w_2 \cdot \mathbb{1}_{\varphi \in J_i} = w_2 \cdot 1 = \frac{1}{2}$.

[1]Although Dietrich (2014) works with complete individual judgments, his definitions naturally extend to the incomplete case.

[2]Note that the term "scoring rules" traditionally refers to a separate class of rules in voting, which we will consider in Section 3.3.

Having a profile $\boldsymbol{J} = (J_1, \dots, J_n)$ and a weight function w, we define the *weight* of judgment J with respect to $\boldsymbol{J}$ by taking the sum over all agents' judgments:

$$w_{\boldsymbol{J}}(J) = \sum_{i \in N} w_{|J_i|}(J)$$

Then, given a superagenda $\Phi^\star$ and a superpopulation $N^\star$, a weight rule F_w determines the collective judgments for any agenda $\Phi \subseteq \Phi^\star$ and group of agents $N \subseteq N^\star$ by selecting the complete and consistent subsets of the agenda with the highest total weight across all agents, for every profile $\boldsymbol{J} = (J_1, \dots, J_n)$:[3][4]

$$F_w(\boldsymbol{J}) \quad = \quad \underset{J \in \mathcal{J}(\Phi)^\bullet}{\text{argmax}}\, w_{\boldsymbol{J}}(J) \quad = \quad \underset{J \in \mathcal{J}(\Phi)^\bullet}{\text{argmax}} \sum_{i \in N} w_{|J_i|} \cdot |J \cap J_i|$$

Given a superagenda $\Phi^\star$, the length of the weight vector $\boldsymbol{w}$ with which a weight rule F_w is associated does not need to exceed the number of all issues in the superagenda, $|\frac{\Phi^\star}{2}|$, since this is the largest possible size of a judgment set that an agent can report. So, (possibly infinite) weight vectors look as follows: $\boldsymbol{w} = (w_\lambda)_{\lambda \in \mathbb{N}, \lambda \leqslant |\Phi^\star|/2}$, with $w_\lambda > 0$ for every λ. For every agenda $\Phi \subseteq \Phi^\star$, only the prefix of $\boldsymbol{w}$ of cardinality $|\frac{\Phi}{2}|$ is relevant for F_w. Furthermore, for every weight rule F_w there are infinitely many weight vectors that induce F_w. For example, if we multiply all weights with some positive constant, then the rule being induced does not change.

In this thesis, a few different weight rules will be spotlighted, due to their intuitive definitions and (as will be shown in Chapter 4) their sensible axiomatic properties. We now turn our attention to them.

3.2. Definition. *Take any weight w with $w_\lambda = w_{\lambda'} = c$ for all $\lambda \leqslant \frac{|\Phi^\star|}{2}$. We call the weight rule F_c that is induced by w the* ***constant weight rule****.*

3.3. Definition. *Take any weight function w with $w_\lambda = \frac{w_1}{\lambda} > 0$ for all $\lambda \leqslant \frac{|\Phi^\star|}{2}$. We call the weight rule F_{ee} that is induced by w the* ***equal-and-even weight rule****.*

This rule, simply induced by the weight vector $\boldsymbol{w} = (1, \frac{1}{2}, \frac{1}{3}, \dots)$, borrows its name from its counterpart for voting with approval preferences, the *equal-and-even cumulative voting rule* (Glasser, 1959; Alcalde-Unzu and Vorsatz, 2009).

Next, we introduce an aggregation rule that makes sense in settings where giving strict priority to agents who report smaller judgment sets is recommendable. In words,

[3]We define weight rules to satisfy *collective rationality*, which requires that all collective judgment sets must be complete and consistent. One could easily relax the completeness assumption, with no important technical implications for our results. In their current form, weight rules produce outcomes that are ready to use: They make a judgment about every issue at stake, without contradictions.

[4]Note that if an agent i does not express a *yes*/*no* judgment on any issue, that is, if $J_i = \emptyset$, then this agent is completely disregarded as far as the collective outcome is concerned. In alternative frameworks, empty judgment sets could influence the collective decision (see, e.g., Slavkovik and Jamroga, 2011).

the *upward-lexicographic* rule first tries to agree as much as possible with the agents holding the smallest judgment sets; in case this process induces more than one collective judgment, the rule aims at maximising agreement with the agents with the second smallest judgment sets, and so on, until no ties can be broken anymore.

In order to proceed to a formal definition of the aggregation rule, we first need an additional technical notion, namely the one of the lexmax function. Consider a set $V = \{\boldsymbol{v}^1, \boldsymbol{v}^2, \ldots, \boldsymbol{v}^m\}$ of equal-sized vectors of real numbers, and denote by $v_i^k \in \mathbb{R}$ the i^{th} element of vector $\boldsymbol{v}^k$. The function lexmax picks up the vector from the set V the elements of which are maximal in a lexicographic sense. More precisely, lexmax : $V \mapsto \boldsymbol{v}^k$, where $\boldsymbol{v}^k \in V$ is such that if $v_i^\ell > v_i^k$ for some $v_i^\ell \in V$, then there exists some $j < i$ for which $v_j^\ell < v_j^k$.

We fix an agenda Φ, a group of agents N, and a profile $\boldsymbol{J} \in \mathcal{J}(\Phi)^n$, and define the "raw points" the agents with judgment sets of size λ assign to choosing J:

$$k_\lambda^{\boldsymbol{J}}(J) = \sum_{i \mid \lambda = |J_i|} |J \cap J_i|$$

Similarly, the "raw points" the agents with judgment sets of size *up to* λ assign to choosing J are captured by the formula

$$K_\lambda^{\boldsymbol{J}}(J) = \sum_{k=1}^{K} k_\lambda^{\boldsymbol{J}}(J)$$

The upward-lexicographic rule selects those complete and consistent judgment sets that result in a lexicographically maximal vector (the argument function arg is defined in the standard way):

$$F_{ulex}(\boldsymbol{J}) = \underset{J \in \mathcal{J}(\Phi)^\bullet}{\text{arglexmax}} \left(K_1^{\boldsymbol{J}}(J), \ldots, K_{|\Phi|/2}^{\boldsymbol{J}}(J) \right)$$

Example 3.4 highlights that applying different weight rules, even on very simple profiles of judgments, can lead to completely diverse outcomes.

3.4. Example. Consider again Example 3.1, where Ann, Bea, and Cal provide the following judgments:

$$J_1 = \{p_1\}$$
$$J_2 = \{\neg p_1, p_2\}$$
$$J_2 = \{\neg p_1, p_2\}$$

Here, propositions p_1 and p_2 are logically independent. Therefore, our weight rules reduce to a simple weighted majority for every proposition separately. If we use the constant weight rule, then the weight of proposition $\neg p_1$ will be twice the weight of proposition p_1, so paper 1 will not be collectively accepted. If we use the equal-and-even weight rule, then p_1 and $\neg p_1$ will get the same weight in total, so the collective decision will be a tie between acceptance and rejection. Then, suppose we apply the upward-lexicographic weight rule. Ann's judgment (J_1) is the smallest one in the profile, so the rule will agree with it, accepting p_1. As far as paper 2 is concerned, since all agents that report a judgment on it agree that it should be accepted, all weight rules will accept it as well. △

Recall that every weight rule is represented by many different weight vectors. The upward-lexicographic rule is special, in the sense that, if we change the weight vector in such a way that some or all of the ratios $\frac{w_{\lambda-1}}{w_\lambda}$ of consecutive weights increase, then we do not change the corresponding aggregation rule. With Lemma 3.5 at hand, we will then be able to give a second definition of the upward-lexicographic rule.

3.5. Lemma. *For any superagenda $\Phi^\star$ and finite superpopulation $N^\star$, let w be any weight function with $\frac{w_{\lambda-1}}{w_\lambda} > \lambda \cdot (|N^\star| - 1)$ for all $\lambda \in \mathbb{N}$, $\lambda \leqslant \frac{|\Phi^\star|}{2}$. Then F_w is the upward-lexicographic rule F_{ulex}.*

Proof. By definition, the upward-lexicographic rule is induced by a weight function w with $\frac{w_{\lambda-1}}{w_\lambda} > \lambda \cdot (|N^\star| - 1)$ for all $\lambda > 1$ and is such that $s_\lambda > 0$ for all λ. It remains to be shown that all weight functions of this kind in fact define the same aggregation rule. So consider any two weight functions w and w' with $\frac{w_{\lambda-1}}{w_\lambda} > \lambda \cdot (|N^\star| - 1)$ and $\frac{w'_{\lambda-1}}{w'_\lambda} > \lambda \cdot (|N^\star| - 1)$ for all $\lambda > 1$. We need to show that $F_w = F_{w'}$. Consider an arbitrary agenda $\Phi \subseteq \Phi^\star$, $n \leqslant |N^\star|$, and profile $\boldsymbol{J} \in \mathcal{J}(\Phi)^n$, as well as two arbitrary judgment sets $J, J' \in \mathcal{J}(\Phi)^\bullet$. We are done if we can show that $w_{\boldsymbol{J}}(J) \geqslant w_{\boldsymbol{J}}(J')$ holds if and only if $w'_{\boldsymbol{J}}(J) \geqslant w'_{\boldsymbol{J}}(J')$ does. We can rewrite $w_{\boldsymbol{J}}(J)$ as follows:

$$w_{\boldsymbol{J}}(J) \;=\; \sum_{i\in N}\sum_{\varphi\in J} w_{J_i}(\varphi) \;=\; \sum_{i\in N} w_{|J_i|}\cdot|J\cap J_i| \;=\; \sum_{\lambda\geqslant 1} w_\lambda \cdot \sum_{i\,|\,\lambda=|J_i|}|J\cap J_i|$$

In case profile $\boldsymbol{J}$ is such that all agents report judgment sets of the same size λ, our claim clearly holds, as then $w_{\boldsymbol{J}}(J) = w_\lambda \cdot \sum_{i\in N}|J\cap J_i|$, meaning that $w_{\boldsymbol{J}}(J) \geqslant w_{\boldsymbol{J}}(J')$ reduces to $\sum_{i\in N}|J\cap J_i| \geqslant \sum_{i\in N}|J'\cap J_i|$ (because we know that $w_\lambda > 0$). So, without loss of generality, we may assume that at most $|N^\star|-1$ agents report a judgment set of the same size and thus $\sum_{i\,|\,\lambda=|J_i|}|J\cap J_i| \leqslant (|N^\star|-1)\cdot\lambda$ for all λ (and accordingly for J'). Together with our assumptions regarding the lower bounds for $\frac{w_{\lambda-1}}{w_\lambda}$ and $\frac{w'_{\lambda-1}}{w'_\lambda}$, this implies that $w_{\boldsymbol{J}}(J) \geqslant w_{\boldsymbol{J}}(J')$ holds if and only if the $\frac{|\Phi|}{2}$-vector $(\sum_{i\,|\,\lambda=|J_i|}|J\cap J_i|)_{\lambda\geqslant 1}$ lexicographically precedes $(\sum_{i\,|\,\lambda=|J_i|}|J'\cap J_i|)_{\lambda\geqslant 1}$. In other words, the relative ranking of J and J' does not depend on the exact choice of weights, provided the condition on ratios is respected, i.e., we have shown that $w_{\boldsymbol{J}}(J) \geqslant w_{\boldsymbol{J}}(J')$ if and only if $w'_{\boldsymbol{J}}(J) \geqslant w'_{\boldsymbol{J}}(J')$. □

We have defined the upward-lexicographic rule in terms of a procedure to be followed to compute the outcome for a given profile. But we can also define the same rule in terms of a weight vector: For $N^\star$ being finite, we observe that, if a weight function w

is such that $\frac{w_{\lambda-1}}{w_\lambda} = \lambda \cdot |N^\star|$, this means that each formula of a small judgment set that appears in a profile (which includes at most $|N^\star|$ agents) has greater value for the aggregation rule F_w than all formulas in larger judgment sets; this precisely captures the intuition behind the upward-lexicographic rule.

3.6. Definition. *Given a superagenda $\Phi^\star$ and a finite superpopulation $N^\star$, take the weight function w such that $w_\lambda = \prod_{k=1}^{\lambda} \frac{1}{k \cdot |N^\star|}$ for all $\lambda \in \mathbb{N}$, $\lambda \leqslant \frac{|\Phi^\star|}{2}$. We call the weight rule F_{ulex} that is induced by w the **upward-lexicographic rule**.*

After having introduced the upward-lexicographic rule, it is natural to define the *downward-lexicographic* rule, which selects those complete and consistent subsets of the agenda that result in a lexicographically maximal vector prioritising the points collected by all judgment sets in a profile. This rule breaks ties by excluding the points collected by individual judgment sets of a fixed size in rounds, moving from the largest to the smallest ones:[5]

$$F_{dlex}(\boldsymbol{J}) = \underset{J \in \mathcal{J}(\Phi)^\bullet}{\text{arglexmax}} \left(K^{\boldsymbol{J}}_{|\Phi|/2}(J), \ldots, K^{\boldsymbol{J}}_1(J) \right)$$

Even though at first sight the upward-lexicographic rule and the downward-lexicographic rule seem to satisfy some notion of duality, we will show that this is not true in general. In particular, for a superagenda that contains all propositional variables, the downward-lexicographic rule is not even a weight rule.

3.7. Proposition. *For any superagenda $\Phi^\star$ that contains all propositional variables and for any superpopulation $N^\star$, the downward-lexicographic rule F_{dlex} is not in the class of weight rules.*

Proof. For a superagenda $\Phi^\star$ that contains all propositional variables and for a superpopulation $N^\star$, we assume, aiming for a contradiction, that the downward-lexicographic rule F_{dlex} is a weight rule. In other words, we will assume that F_{dlex} is induced by a weight w, or equivalently, by a weight vector $\boldsymbol{w} = (w_1, w_2, \ldots)$. We show two things:

(i) The following must hold:

$$w_\lambda > w_{\lambda+1} > 0, \text{ for all } \lambda \in \mathbb{N}$$

This means that $\boldsymbol{w}$ corresponds to a decreasing and bounded sequence of real numbers (note that the weights are strictly positive by definition), which by the monotone convergence theorem (see, e.g., Schechter, 1996) has to converge to a non-negative real number.

[5] The definition of the upward-lexicographic rule is not affected if we substitute all $K^{\boldsymbol{J}}_\lambda$ by $k^{\boldsymbol{J}}_\lambda$, while for the downward-lexicographic rule it is (and the new rule we obtain is not intuitively attractive).

(ii) The following must also hold:

$$w_\lambda - w_{\lambda+1} > w_{\lambda-1} - w_\lambda, \text{ for all } \lambda > 1$$

In plain words, consecutive members of the sequence $\boldsymbol{w}$ must be further and further from each other as λ grows large, and hence the sequence cannot converge.

If we know that conditions (i) and (ii) are true, we will have that the sequence of weights in $\boldsymbol{w}$ converges. Take $0 < \alpha < w_{\lambda-1} - w_\lambda$. By the definition of convergence, there must exist some $k' \in \mathbb{N}$ such that $w_k - w_{k+1} < \alpha$ for all $k \geqslant k'$. But by (ii), this is impossible. So a contradiction is reached, and we can conclude that the downward-lexicographic rule F_{dlex} is not in the class of weight rules.

It remains to prove the statements (i) and (ii). Consider an arbitrary $\lambda > 0$ and a consistent judgment set J of size λ such that $J' = (J \setminus \{\varphi\}) \cup \{\neg\varphi\}$ is also consistent, where $\varphi \in J$ (such a J exists because the superagenda contains all propositional variables; for example, J can be constructed to only contain atomic formulas). Take the agenda Φ that includes all issues appearing in J.

For (i), take J and J' as just described, and such that $|J| = |J'| = \lambda + 1$. Then, consider a judgment set $J'' = J' \setminus \{\psi\}$, where $\psi \neq \neg\varphi$ (that is, $|J''| = \lambda$). Using the profile $\boldsymbol{J}' = (\emptyset, \dots, \emptyset, J, J'')$ for an arbitrary group $N \subseteq N^\star$, the downward-lexicographic rule prescribes that $F_{dlex}(\boldsymbol{J}') = \{J'\}$, which implies that $w_\lambda > w_{\lambda+1}$.

For (ii), assume that $\lambda > 1$ and consider the judgment set $J''' = J \setminus \{\psi', \psi''\}$, where $\psi', \psi'' \neq \varphi$, which is of size $\lambda - 1$. We construct the following profile, again for an arbitrary group $N \subseteq N^\star$:

$$\boldsymbol{J}''' = (\emptyset, \dots, \emptyset, J'', J'', J, J''')$$

$$\begin{aligned} \text{Then,} \quad & K^{\boldsymbol{J}'''}_{\lambda+1}(J) = 2 \cdot (\lambda - 1) + \lambda + 1 + \lambda - 1 = 2 \cdot \lambda + \lambda + \lambda - 2 = K^{\boldsymbol{J}'''}_{\lambda+1}(J') \\ \text{and} \quad & K^{\boldsymbol{J}'''}_{\lambda}(J) = 2 \cdot (\lambda - 1) + \lambda - 1 < 2 \cdot \lambda + \lambda - 2 = K^{\boldsymbol{J}'''}_{\lambda}(J'), \end{aligned}$$

which means that $F_{dlex}(\boldsymbol{J}''') = \{J'\}$. Thus it must be the case that $2 \cdot w_\lambda > w_{\lambda-1} + w_{\lambda+1}$. Equivalently: $w_\lambda - w_{\lambda+1} > w_{\lambda-1} - w_\lambda$. □

Finally, let us remind the reader of the Kemeny rule, applicable in situations where all agents submit complete judgment sets. The class of weight rules can be seen as a generalisation of Kemeny for incomplete individual judgments. This is straightforward, since the weights vary in the size of the reported judgment sets only—when all these sets are complete, the weights do not play any role.

3.8. Proposition. *For every profile $\boldsymbol{J} \in \mathcal{J}(\Phi)^\bullet$ of complete individual judgments, and for every weight function w, it holds that $F_w(\boldsymbol{J}) = Kemeny(\boldsymbol{J})$.*

3.1.2 Voting and Preference Aggregation

In this section, we transfer the concept of weight rules to the frameworks of voting and preference aggregation. We thus zoom on a family of aggregation rules under which the *weight* assigned to agent i's ranking of two alternatives depends on how many other pairs of alternatives she ranks as well.

3.9. Example. Suppose that Ann, Bea, and Cal are asked to compare three hotels (a, b, and c) according to their preferences. Every agent holds preferences about the hotels she has visited and only those, giving us the following profile:

$$\begin{tikzcd}[column sep=small]
 & \text{Ann} & & & & \text{Bea} & & & & \text{Cal} & \\
a \arrow[rr] \arrow[dr] & & c & & a \arrow[rr] & & c & & a & & c \arrow[dl] \\
 & b & & & & b \arrow[ul] \arrow[ur] & & & & b &
\end{tikzcd}$$

Ann and Bea, for example, both express a ranking between alternatives a and b, but they disagree on it; at the same time, Bea reports one more pairwise comparison than Ann. How should a weight rule deal with this information? $\triangle$

We consider different variants of weight rules, concerning the *type* of output required: We may use a rule to obtain a consensus regarding the *ranking* of the alternatives, we may look for a single *winner* amongst the alternatives, or we may need to select a *set of winners* of a given size. Our model is flexible with respect to the collective outcomes it produces. As far as the input of the aggregation is concerned, i.e., the type of the individual preferences, we study strict and acyclic preferences (preferences P in the domain $\mathcal{P}(A)$). Working with strict preferences facilitates the utilisation of the relevant framework we have developed for judgment aggregation in Section 3.1.1 above, since every ranked pair of alternatives ab directly corresponds to a proposition "a is better than b". Still, as we will see, ample new questions arise in the context of preference aggregation, and original insights are required.

The primary constituent of a *weight rule* for aggregating preferences is again a *weight function* $w : \mathbb{N} \to \mathbb{R}^+$, as we defined for judgment aggregation. A weight function assigns to each natural number, representing a preference size $|P_i|$, a positive real number $w_{|P_i|} = w(|P_i|)$, expressing how much any ordered pair $ab \in P_i$ weighs, given the size of P_i.

We use the following notation to refer to the total weight assigned to all the pairs of alternatives in a preference $P \subseteq A{\times}A$ from the perspective of holding P_i:

$$w_{P_i}(P) \;=\; \sum_{ab\in P} w_{|P_i|} \cdot \mathbb{1}_{ab\in P_i} \;=\; w_{|P_i|} \cdot |P \cap P_i|$$

Another component of the definition of a weight rule for preference aggregation is a *type* $\mathcal{T}$, which is a function that maps every set of alternatives $A \subseteq A^\star$ to a subset $\mathcal{T}(A) \subseteq \mathcal{P}(A)$ of preference sets. We examine four specific types.[6]

- First, in the literature on preference aggregation both individual preferences and outcomes traditionally are *complete rankings* over the alternatives. So we consider the type of *linear orders* $\mathcal{L}$, where $\mathcal{L}(A)$ contains all acyclic, strict, and complete, relations over A.

- Second, in voting theory we are mainly concerned with finding a (single) *winner*, that is, an alternative ranked higher than any other alternative (Zwicker, 2016). Hence, we also define the type $\mathcal{W}$, where $\mathcal{W}(A)$ includes all pairwise preferences of the form $P = \{aa' \mid a' \in A \setminus \{a\}\}$, for some $a \in A$.

- Third, in multiwinner voting the outcome should be a given number of winning alternatives, constituting, for instance, a *committee* (Faliszewski et al., 2017). So we define the type $\mathcal{W}_k$, where $\mathcal{W}_k(A)$ consists of all pairwise preferences that designate a set of k winners (for some fixed $k > 1$), i.e., that are of the form $P = \{(a_1, a'), \ldots, (a_k, a') \mid a' \in A \setminus \{a_1, \ldots, a_k\}\}$, for some $a_1, \ldots, a_k \in A$ with $a_i \neq a_j$ for $i \neq j$.

- Fourth, we define the type C of *chains*, which are linear orders over subsets of A, to model situations where agents rank part of the alternatives. Formally, $C(A)$ consists of all pairwise preferences of the form $P = (a_1 \rhd a_2 \rhd \cdots \rhd a_k)$, for some $k \in \mathbb{N}$ and $a_1, \ldots, a_k \in A$ with $a_i \neq a_j$ for $i \neq j$.

We use the letter $\mathcal{P}$ to denote the type of *all* preference sets in $\mathcal{P}(A)$. Given a weight function w, a *weight rule of type* $\mathcal{T}$, denoted by $F_w^{\mathcal{T}}$, decides the outcome of a profile $\boldsymbol{P} = (P_1, \ldots, P_n)$, for every specific set of alternatives $A \subseteq A^\star$ and group of agents $N \subseteq N^\star$, by selecting those preference sets in $\mathcal{T}(A)$ that maximise the total weight across all agents:

$$F_w^{\mathcal{T}}(\boldsymbol{P}) \quad = \quad \underset{P \in \mathcal{T}(A)}{\mathrm{argmax}} \sum_{i \in N} w_{P_i}(P)$$

As we pointed out for weight rules in judgment aggregation too, multiplying all weights with any positive real number will not affect the induced rule. Two weight rules for preference aggregation will play a role in this thesis.

3.10. Definition. *Take any weight function w with $w_\lambda = w_{\lambda'}$ for all $\lambda, \lambda' \in \mathbb{N}$. We call the weight rule $F_c^{\mathcal{T}}$ induced by w the **constant weight rule** of type $\mathcal{T}$.*

[6]Lang et al. (2017) define certain aggregation outcomes that include ours. For instance, their *plain dominating 1-chain* and *plain dominating k-subset* correspond to our single-winner and k-winner sets, respectively. But the model and the results they provide are restricted to complete input profiles.

3.11. Definition. *Take any weight function w with $w_\lambda = \frac{w_1}{\lambda}$ for all $\lambda \in \mathbb{N}$. We call the weight rule $F_{ee}^{\mathcal{T}}$ induced by w the* ***equal-and-even weight rule*** *of type $\mathcal{T}$.*

These rules are natural analogues of the ones of Section 3.1.1, in judgment aggregation (note that we do not include an upward-lexicographic rule in our analysis for preference aggregation, although such a rule could also be constructed). Let us now try to better understand how they work, returning to Example 3.9. Without loss of generality, we set w_1 to be equal to 1.

When the constant weight F_c rule is applied, every pairwise comparison reported by the agents weighs the same. For outcomes that are linear orders, the rule will try to agree as much as possible with the input provided by the agents, similarly to the standard Kemeny rule in preference aggregation. The collective preferences will then be $a \rhd b \rhd c$, $a \rhd c \rhd b$, and $b \rhd a \rhd c$, tied, with weight 4. For single-winner outcomes, the preference corresponding to a being the winner will be the collective decision with weight 3, and if we seek two winners, then a and b will do the job.

When the equal-and-even weight rule is applied, then every pairwise comparison will be weighed as follows:

Ann

$$\begin{tikzcd} a \arrow[rr, "1/2"] \arrow[dr, "1/2"'] & & c \\ & b & \end{tikzcd}$$

Bea

$$\begin{tikzcd} a \arrow[rr, "1/3"] & & c \\ & b \arrow[ul, "1/3"] \arrow[ur, "1/3"'] & \end{tikzcd}$$

Cal

$$\begin{tikzcd} a & & c \arrow[dl, "1"] \\ & b & \end{tikzcd}$$

For outcomes that are linear orders, the collective preference now is $a \rhd c \rhd b$, with weight $\frac{1}{3} + \frac{1}{2} + \frac{1}{2} + 1$. For outcomes of single-winner type, the winner is again a, with the corresponding preference getting weight $\frac{1}{3} + \frac{1}{2} + \frac{1}{2}$.

Some further observations can be made regarding the constant weight rule and the equal and even weight rule, and the way they operate for different types of outputs, when the input consists of profiles of chains (i.e., on profiles $\boldsymbol{C} \in \mathcal{C}(A)^n$).

We first inspect the constant, winner-type weight rule on chains, and find that it coincides with a variant of the Borda rule: For a chain $a_1 \rhd a_2 \rhd \cdots \rhd a_k$, a_1 gets $k-1$ points, a_2 gets $k-2$ points, and so forth, with a_k and all alternatives not appearing in the chain getting 0 points. *Shortsighted Borda (sBorda)* maximises the collected points in a profile (Terzopoulou and Endriss, 2019a). This version of Borda agrees with the rule assigning a score to an alternative based on the length of the *longest path* below it (Endriss et al., 2009).

The constant weight k-winner rule is not k-sBorda for profiles of chains,[7] but it gives rise to a neat original definition of a multiwinner rule. Conceptually, when deciding for a set of k winners, one could regard the potential outcome (for instance, the committee)

[7]For a counterexample, consider a profile with three preferences that are chains: $a \rhd b \rhd c \rhd d$, $c \rhd d \rhd a$, and $d \rhd a$.

as a whole and compare the power of its members (based on the voters' explicitly declared preferences) against its nonmembers: Consider the multiwinner voting rule that makes a set $S \subseteq A$ of fixed size win when the alternatives in S are preferred over the alternatives outside S the maximum number of times.

$F^{\mathcal{W}}_{ee}$, when applied to profiles of chains, is reminiscent of a positional scoring rule, suggesting the definition of a new family of voting rules that integrates positional scoring rules (Zwicker, 2016) and cumulative voting rules (Glasser, 1959): Consider the rule that associates a chain $a_1 P a_2 P \cdots P a_{k-1} P a_k$ of length k with the weight vector $(\frac{k-1}{k(k-1)/2}, \frac{k-2}{k(k-1)/2}, \dots, \frac{1}{k(k-1)/2}, 0)$. Under this rule, every agent distributes a total weight of 1 across the alternatives in her chain in a Borda-like fashion.

The equal-and-even weight k-winner rule, still for profiles of chains, closely follows the constant weight rule of the same type; however, in this case, the power of a committee member over an alternative outside the committee hinges also on the sizes of the preferences that rank the member above the nonmember.

Type Restrictions

The *restriction* of a set of preferences $S \subseteq \mathcal{P}(A)$ to a given type $\mathcal{T}$ is formally defined as $S|_{\mathcal{T}} = S \cap \mathcal{T}(A)$.

Recall that, to compute $F^{\mathcal{T}}_w(\boldsymbol{P})$, we go through all preferences P of type $\mathcal{T}$ and select those that maximise total weight. An alternative definition would be to first select all preferences P of *any* type that maximise total weight and to then restrict the resulting set to $\mathcal{T}$, i.e., to compute $(\mathrm{argmax}_{P \in \mathcal{P}(A)} \sum_{i \in N} w_{P_i}(P))|_{\mathcal{T}}$, which we can also write as $F^{\mathcal{P}}_w(\boldsymbol{P})|_{\mathcal{T}}$. When do these two alternative approaches yield the same outcomes?

To answer this question, we require some further terminology. Given two types $\mathcal{T}$ and $\mathcal{T}'$ we say that $\mathcal{T}$ *refines* $\mathcal{T}'$ if $\mathcal{T}(A) \subseteq \mathcal{T}'(A)$ for all $A \subseteq A^\star$. We furthermore say that $\mathcal{T}$ *extends* $\mathcal{T}'$ if for all $A \subseteq A^\star$ and preferences $P' \in \mathcal{T}'(A)$ there exists a preference $P \in \mathcal{T}(A)$ with $P \supseteq P'$.

3.12. Proposition. *If $\mathcal{T}$ simultaneously refines and extends $\mathcal{T}'$, then $F^{\mathcal{T}}_w(\boldsymbol{P}) = F^{\mathcal{T}'}_w(\boldsymbol{P})|_{\mathcal{T}}$ for any w and $\boldsymbol{P}$.*

Proof. Since $\mathcal{T}(A) \subseteq \mathcal{T}'(A)$, which holds as $\mathcal{T}$ refines $\mathcal{T}'$, we have that $F^{\mathcal{T}}_w(\boldsymbol{P}) = \mathrm{argmax}_{P \in \mathcal{T}(A)} \sum_{i \in N} w_{P_i}(P) \supseteq \mathcal{T}(A) \cap \mathrm{argmax}_{P \in \mathcal{T}'(A)} \sum_{i \in N} w_{P_i}(P) = F^{\mathcal{T}'}_w(\boldsymbol{P})|_{\mathcal{T}}$.

For the other direction, we first show that there exists at least one $P \in \mathcal{T}(A)$ with $P \in F^{\mathcal{T}'}_w(\boldsymbol{P})$. Take any $P' \in F^{\mathcal{T}'}_w(\boldsymbol{P})$. Since $\mathcal{T}$ extends $\mathcal{T}'$, there exists an $R \in \mathcal{T}(A)$ with $R \supseteq R'$. Observe the following:

$$w_{|P_i|} \cdot |P_i \cap P| \geqslant w_{|P_i|} \cdot |P_i \cap P'| \text{ for all } i \in N$$

Thus, the following inequality also holds:

$$\sum_{i \in N} w_{P_i}(P) \geqslant \sum_{i \in N} w_{P_i}(P')$$

Hence, we indeed get $P \in F_w^{\mathcal{T}'}(\boldsymbol{P})$ as claimed.

Now take any $P^* \in F_w^{\mathcal{T}}(\boldsymbol{P})$. By definition of the weight rule, this entails that the following holds, including the specific R we know to exist in $F_w^{\mathcal{T}'}(\boldsymbol{P})$:

$$\sum_{i \in N} w_{P_i}(P^*) \geqslant \sum_{i \in N} w_{P_i}(P) \text{ for all } R \in \mathcal{T}(A)$$

As furthermore $P^* \in \mathcal{T}'(A)$, which is due to $\mathcal{T}(A) \subseteq \mathcal{T}'(A)$, we get $P^* \in F_w^{\mathcal{T}'}(\boldsymbol{P})$ and thus also $P^* \in F_w^{\mathcal{T}'}(\boldsymbol{P})|_{\mathcal{T}}$. We conclude that $F_w^{\mathcal{T}}(\boldsymbol{P}) \subseteq F_w^{\mathcal{T}'}(\boldsymbol{P})|_{\mathcal{T}}$. □

Every type $\mathcal{T}$ refines $\mathcal{P}$. But which types $\mathcal{T}$ also extend $\mathcal{P}$? Certainly $\mathcal{L}$ extends $\mathcal{P}$, as every pairwise-preference set can be extended to a linear order, and therefore so does every type $\mathcal{T}$ that includes all linear orders, answering our question.

3.13. Corollary. *For any weight function w and type $\mathcal{T}$ that is refined by $\mathcal{L}$, it holds that $F_w^{\mathcal{T}}(\boldsymbol{P}) = F_w^{\mathcal{P}}(\boldsymbol{P})|_{\mathcal{T}}$.*

Weight rules reduce to some known rules in the special case of complete preference profiles: The Kemeny rule in the context of preference aggregation (that is, for outcomes that are linear order), the Borda rule in single-winner voting, and the k-Borda rule in multiwinner voting with k winners.

3.14. Proposition. *For all $\boldsymbol{L} \in \mathcal{L}(A)^n$ and weight functions w, the following hold:*

- $F_w^{\mathcal{L}}(\boldsymbol{L}) = \mathit{Kemeny}(\boldsymbol{L})$
- $F_w^{\mathcal{W}}(\boldsymbol{L}) = \mathit{Borda}(\boldsymbol{L})$
- $F_w^{\mathcal{W}_k}(\boldsymbol{L}) = k\text{-}\mathit{Borda}(\boldsymbol{L})$

Proof. The statement clearly holds for the first case, where the output of the aggregation is of the type of linear orders: When all input preferences are complete, the weights corresponding to the sizes of these preferences are not important, since they all boil down to the same constant number—the definition of every weight rule then coincides with the definition of the Kemeny rule.

For the second case, where the output distinguishes single winners, the statement follows from a careful observation regarding the scores associated with the Borda rule. Consider a preference $P^a \in \mathcal{W}(A)$ that has alternative a on top. This preference will be in the outcome of $F_w^{\mathcal{W}}(\boldsymbol{L})$, declaring a as the winner, if it gets maximum total weight by the rule $F_w^{\mathcal{W}}$ in the profile $\boldsymbol{L}$. The weight assigned to P^a from every preference $L \in \mathcal{L}(A)$ is determined by how many ordered pairs of alternatives R^a and L have in common—since R^a contains exactly the pairs where alternative a is considered better than another alternative b for all $b \in A \setminus \{a\}$, its relevant weight from L is exactly equal to the position of the alternative a in the (transitive) ranking of L.[8]

[8] A similar result is also obtained in the work of Lang et al. (2017), and in that of Endriss (2018), who simulates voting in the framework of judgment aggregation.

The reasoning for the third case with winning sets of alternatives is the same as the single-winner case above. □

Comparing to the Concepts of Possible and Necessary Winners

A highly influential concept regarding the aggregation of incomplete preferences was developed by Konczak and Lang (2005) and was subsequently studied further by several authors (see, e.g., Pini et al., 2007; Xia and Conitzer, 2008; Betzler and Dorn, 2010). The idea is that, given an aggregation rule F that is defined only on complete inputs and a profile $\boldsymbol{P}$ with incomplete preferences, we can consider all those profiles that contain completions of the preferences in $\boldsymbol{P}$ and apply F on them. We say that P_i^* is a *completion* of P_i if $P_i^* \in \mathcal{L}(A)$ and $P_i^* \supseteq P_i$.

Then, an outcome is *possible* (*necessary*) if it is an outcome for some (all) completion(s) of $\boldsymbol{P}$, and an alternative is a *possible* (*necessary*) *winner* if it is the top alternative of a possible (necessary) outcome. We denote by $P^r(F, \boldsymbol{P})$, $P^w(F, \boldsymbol{P})$, and $N^r(F, \boldsymbol{P})$, $N^w(F, \boldsymbol{P})$, the sets of possible and necessary rankings and winners, respectively.

These concepts are relevant in situations where the incompleteness of the agents' preferences is due to the lack of information we have about those preferences. In this thesis, we instead study intrinsically incomplete preferences. Nevertheless, it is tempting to conjecture that our rules calculate (part of) the possible outcomes under the terminology of Konczak and Lang. For instance, the possible outcomes of the Kemeny rule might get confused with the outcomes induced by the constant weight rule $F_c^{\mathcal{L}}$. However, this is not true. We find that the outcomes of the constant weight rule in fact can be totally distinct from both the possible and the necessary outcomes under the Kemeny rule. Given a set S of linear orders, we write *top-alt*(S) for those alternatives that appear in the top position of some order in S.

3.15. Proposition. *There exist (different) profiles $\boldsymbol{P} \in \mathcal{P}(A)^n$ and (different) alternatives $a \in A$ such that the following hold:*

(1) $a \in P^w(\textit{Kemeny}, \boldsymbol{P})$, $a \in N^w(\textit{Kemeny}, \boldsymbol{P})$, $a \notin \textit{top-alt}(F_c^{\mathcal{L}}(\boldsymbol{P}))$

(2) $a \notin P^w(\textit{Kemeny}, \boldsymbol{P})$, $a \notin N^w(\textit{Kemeny}, \boldsymbol{P})$, $a \in \textit{top-alt}(F_c^{\mathcal{L}}(\boldsymbol{P}))$

(3) $a \in P^w(\textit{Kemeny}, \boldsymbol{P})$, $a \notin N^w(\textit{Kemeny}, \boldsymbol{P})$, $a \notin \textit{top-alt}(F_c^{\mathcal{L}}(\boldsymbol{P}))$

(4) $a \in P^w(\textit{Kemeny}, \boldsymbol{P})$, $a \notin N^w(\textit{Kemeny}, \boldsymbol{P})$, $a \in \textit{top-alt}(F_c^{\mathcal{L}}(\boldsymbol{P}))$

(5) $a \in P^w(\textit{Kemeny}, \boldsymbol{P})$, $a \in N^w(\textit{Kemeny}, \boldsymbol{P})$, $a \in \textit{top-alt}(F_c^{\mathcal{L}}(\boldsymbol{P}))$

And this is an exhaustive list of all possible cases.

Proof. For statement (1): Take $A = \{a, b, c\} \subseteq A^\star$, and consider the profile $\boldsymbol{P} = (\{cb, ba, ca\}, \{ac, cb\})$ and its completions in $\mathcal{L}(A)^2$:

$$\begin{array}{l|l} P_1 = \{cb, ba, ca\} & c \rhd b \rhd a \\ P_2 = \{ac, cb\} & a \rhd c \rhd b \end{array}$$

Then, $a \in P^w(\mathit{Kemeny}, \boldsymbol{P}) = N^w(\mathit{Kemeny}, \boldsymbol{P}) = \{a, c\}$, but $a \notin \mathit{top\text{-}alt}(F_c^{\mathcal{L}}(\boldsymbol{P})) = \{c\}$.

For statement (2): Take $A = \{a, b, c\} \subseteq A^\star$, and consider the profile $\boldsymbol{P} = (\{ab, bc, ac\}, \{ab, ca\}, \{ab, ca\})$ and its completions in $\mathcal{L}(A)^2$:

$$\begin{array}{l|l} P_1 = \{ab, bc, ac\} & a \rhd b \rhd c \\ P_2 = \{ca, ab\} & c \rhd a \rhd b \\ P_3 = \{ca, ab\} & c \rhd a \rhd b \end{array}$$

So, $a \notin P^w(\mathit{Kemeny}, \boldsymbol{P}) = N^w(\mathit{Kemeny}, \boldsymbol{P}) = \{c\}$, but $a \in \mathit{top\text{-}alt}(F_c^{\mathcal{L}}(\boldsymbol{P})) = \{a, c\}$.

For statement (3): Take $A = \{a, b, c\} \subseteq A^\star$, the profile $\boldsymbol{P} = (P_1, P_2) \in \mathcal{P}(A)^2$, and its three completions in $\mathcal{L}(A)^2$:

$$\begin{array}{l|lll} P_1 = \{cb, ba, ca\} & c \rhd b \rhd a & c \rhd b \rhd a & c \rhd b \rhd a \\ P_2 = \{ac\} & b \rhd a \rhd c & a \rhd b \rhd c & a \rhd c \rhd b \end{array}$$

Then, $a \in P^w(\mathit{Kemeny}, \boldsymbol{P})$, yet $N^w(\mathit{Kemeny}, \boldsymbol{P}) = \{c\} = \mathit{top\text{-}alt}(F_c^{\mathcal{L}}(\boldsymbol{P}))$.

For statement (4): Take $A = \{a, b, c\} \subseteq A^\star$, and consider the profile $\boldsymbol{P} = (\{ab, bc, ac\}, \{ca, ab\}, \{ca\})$ and its three completions in $\mathcal{L}(A)^2$:

$$\begin{array}{l|lll} P_1 = \{ab, bc, ac\} & a \rhd b \rhd c & a \rhd b \rhd c & a \rhd b \rhd c \\ P_2 = \{ca, ab\} & c \rhd a \rhd b & c \rhd a \rhd b & c \rhd a \rhd b \\ P_3 = \{ca\} & c \rhd a \rhd b & b \rhd c \rhd a & c \rhd b \rhd a \end{array}$$

So, $a \in P^w(\mathit{Kemeny}, \boldsymbol{P}) = \mathit{top\text{-}alt}(F_c^{\mathcal{L}}(\boldsymbol{P})) = \{a, b, c\}$, but $a \notin N^w(\mathit{Kemeny}, \boldsymbol{P}) = \{c\}$.

For statement (5): Take $A = \{a, b, c\} \subseteq A^\star$, and consider the profile $\boldsymbol{P} = (\{ab, bc, ac\}, \{ca, ab\}, \{ca\})$ and its three completions in $\mathcal{L}(A)^2$:

$$\begin{array}{l|lll} P_1 = \{cb, ba, ca\} & c \rhd b \rhd a & c \rhd b \rhd a & c \rhd b \rhd a \\ P_2 = \{ac, cb\} & a \rhd c \rhd b & a \rhd c \rhd b & a \rhd c \rhd b \\ P_3 = \{ac\} & a \rhd c \rhd b & b \rhd a \rhd c & a \rhd b \rhd c \end{array}$$

So, $a \in P^w(\mathit{Kemeny}, \boldsymbol{P}) = \mathit{top}(F_c^{\mathcal{L}}(\boldsymbol{P})) = \{a, b, c\}$, and $a \in N^w(\mathit{Kemeny}, \boldsymbol{P}) = \{a\}$.

We have thus finished with all cases. □

Going beyond constant weights, we also show that, independently of the weight function w, a necessary Kemeny ranking does not have to be an outcome under $F_w^{\mathcal{L}}$. This refutes the intuition that necessary outcomes would be included in the outcomes of (at least some) weight rules.

3.16. Proposition. *There exists a profile $\boldsymbol{P} \in \mathcal{P}(A)^n$ such that for any weight rule $F_w^{\mathcal{L}}$, $N^r(\mathit{Kemeny}, \boldsymbol{P}) \not\subseteq F_w^{\mathcal{L}}(\boldsymbol{P})$.*

Proof. Consider the set $A = \{a, b, c, d\}$ and the profile $\boldsymbol{P} = (\{ab, bc, ac, ad, dc\}, \{ab, ca, da, dc\})$ with two agents. By the calculations, we have that $N^r(\mathit{Kemeny}, \boldsymbol{P}) = \{L \in \mathcal{L}(A) \mid L \supseteq \{ab, dc, db\}\}$. Now suppose, aiming for a contradiction, that $N^r(\mathit{Kemeny}, \boldsymbol{P}) \subseteq F_w^{\mathcal{L}}(\boldsymbol{P})$ for some weight rule $F_w^{\mathcal{L}}(\boldsymbol{P})$. This would mean that the preference sets $\{ab, bc, ac, da, dc\}$, $\{ab, bc, ac, ad, dc\}$, and $\{ab, ca, da, dc\}$ (the completions of which belong to the necessary outcomes) have the same total weight in $\boldsymbol{P}$. Thus, it would be the case that $4w_5 + 3w_4 = 5w_5 + 2w_4 = 2w_5 + 4w_4 \Leftrightarrow w_4 = w_5 = 0$, contradicting our assumption that $w_\lambda > 0$ for every λ. □

> Allowing for non-transitive individual preferences is essential for the construction of the profiles used in Propositions 3.15 and 3.16.

Finally, let us stress that there is also a clear difference in complexity between using weight rules and computing possible winners: $F_w^{\mathcal{W}}$ generalises the Borda rule, and its outcomes can be computed in polynomial time. Yet, the possible-winner determination problem for Borda is NP-complete (Xia and Conitzer, 2008).

3.2 Quota Rules in Judgment Aggregation

This section is centered on the framework of judgment aggregation, and introduces the class of quota rules. Quota rules are very natural aggregation rules, requiring that a proposition be collectively accepted if and only if the number of agents that agree with it exceeds a given threshold. They provide a simple—easy to compute and easy to explain—method for aggregating judgments that is commonly used in practice. For instance, the referendum about monarchy/parliamentarism in Brazil[9] and the post-war referendum in Italy[10] both concerned correlated issues and adopted quota rules.

When all agents are expected to submit complete opinions, quota rules are defined in a straightforward fashion. But in cases where the agents may also abstain on some of the issues and report incomplete judgments, there are several ways for determining the relevant threshold, depending on the number of abstentions or the margin between those that agree and those that disagree with a given proposition. We here systematically design quota rules for incomplete inputs. Let us motivate our work with an example.

3.17. Example. Consider the board of a college in a small town, having to decide whether to offer fresh fruit on campus during the coming academic year. In the town

[9] https://web.archive.org/web/20200501204603/https://en.wikipedia.org/wiki/1993_Brazilian_constitutional_referendum

[10] https://web.archive.org/web/20201130140318/https://en.wikipedia.org/wiki/1946_Italian_general_election

there are only two farms (farm A and farm B) that can supply the college with fresh fruit. The board members are thus asked to express their judgments on three issues: whether a contract with farm A should be established, whether a contract with farm B should be established, and whether fresh fruit should be offered on campus. It happens that an 80% supermajority of the board members do not have any particular opinion about fruit, or about farm A, but they really dislike farm B (which is rumoured to follow unethical animal treatments). The rest of the board members report a clear judgment in favour of farm B and against farm A, while also supporting the offer of fruit on campus.

	FARM A	FARM B	FRUIT ON CAMPUS
80%	☐	☒	☐
20%	☒	☑	☑

How should the board decide? On the one hand, all members that expressed *some* opinion about farm A were negative regarding a contract with it, and all those that did not abstain with respect to the fruit issue were positive towards it. It would then be natural for the board to respect these unanimous opinions. On the other hand, a straightforward majority of the members were opposed to a contract with farm B, leading to an impossible situation for the board: it would need to be able to offer fresh fruit on campus without establishing a contract with any of the two providers. $\triangle$

The way to count agents who abstain is not unequivocal: When designing an aggregation rule, we may want to only rely on those agents that actually report a positive or a negative judgment, or we may also want to take into consideration those that abstain. What option is better varies with the context. In practice, incomplete judgments are (or should be) treated in several different ways, depending on the specific situation at hand and the institution where the decision making takes place. We will present four alternative formulations of quota rules that capture different applications and study how they relate to each other.[11] We will see that one of these formulations is the most general of all, hinging on thresholds that vary in the number of abstentions.

Formal Definitions of Quota Rules for Incomplete Judgments. The threshold of acceptance for a proposition φ is defined based on (i) the *absolute number* of φ's supporters, or (ii) the *margin* of those who support φ over those who support $\neg\varphi$. We consider two versions of each one of the aforementioned cases, regarding whether or not the relevant threshold depends on the total number of abstentions on $\widetilde{\varphi}$: if it does, the threshold is called *variable*, otherwise it is called *invariable*. We thus have four classes of quota rules, operating for every agenda $\Phi \subseteq \Phi^\star$ and group $N \subseteq N^\star$ (below, we present them for a fixed agenda Φ and group N).

Table 3.1 provides a taxonomy of the different possibilities for defining quota rules for incomplete judgments, upon which we will elaborate soon.

[11]Note, though, that we do not distinguish between agents who do not go to vote and those who submit a blank/invalid vote, although many aggregation mechanisms in practice do.

Abstentions \ Margins	do not matter	matter
do not matter	invariable absolute	invariable marginal
matter	variable absolute	variable marginal

Table 3.1: A taxonomy of quota rules for incomplete judgments.

The first class contains the rules according to which an absolute threshold has to be reached for a decision to be made, independently of the possible abstentions.

3.18. Definition. *Consider a function $a : \Phi \to \{0, \dots, n+1\}$, and denote the number $a(\varphi)$ by a_φ. The* ***invariable absolute quota rule*** *F_a is such that:*

$$F_a(\boldsymbol{J}) = \{\varphi \in \Phi \mid n_\varphi^{\boldsymbol{J}} \geqslant a_\varphi\}$$

For example, according to Article 27 of the UN Charter, decisions of the United Nations Security Council on procedural matters (φ) are confirmed if and only if there is an affirmative vote of at least nine (out of the fifteen) members of the council. So, an invariable absolute threshold $a_\varphi = 9$ is employed in practice.

The second class contains rules imposing a minimum margin between positive and negative votes, still independently of the possible abstentions.

3.19. Definition. *Consider a function $m : \Phi \to \{-n, \dots, n+1\}$ and denote the number $m(\varphi)$ by m_φ. The* ***invariable marginal quota rule*** *F_m is such that:*

$$F_m(\boldsymbol{J}) = \{\varphi \in \Phi \mid n_\varphi^{\boldsymbol{J}} - n_{\neg\varphi}^{\boldsymbol{J}} \geqslant m_\varphi\}$$

In the Brexit referendum of 2016, the difference between the *Leave* and the *Remain* votes was not even 1.3 million, while the population of the United Kingdom is around 66 millions. An important decision with big consequences was decided by a very small margin (less than 2% of the British people). For such a crucial change to be implemented, many would find it desirable to have a rule requiring a minimum margin between positive and negative votes. Hence, although this was not the case in reality, an invariable marginal quota rule with threshold $m_\varphi = c$ for some large constant c would probably be appropriate.

Next, we have the class of rules that impose a threshold of acceptance for a proposition depending on the total number of reported judgments. In a variable threshold a_φ^k or m_φ^k, the parameter k codifies the total number of reported judgments about $\widetilde{\varphi}$.

3.20. Definition. *Consider a function $va : \{0, \dots, n\} \times \Phi \to \{0, \dots, n+1\}$ and denote the number $va(k, \varphi)$ by a_φ^k. The* ***variable absolute quota rule*** *F_{va} is such that:*

$$F_{va}(\boldsymbol{J}) = \{\varphi \in \Phi \mid n_\varphi^{\boldsymbol{J}} \geqslant a_\varphi^{n_{\widetilde{\varphi}}^{\boldsymbol{J}}}\}$$

The most popular way of conducting a referendum (for instance in Switzerland) is by letting a proposal go through if and only if it is accepted by the majority of the submitted votes.[12] Such referendums in practice use the simple majority rule, with a variable absolute threshold $a^k_\varphi = \lceil \frac{k}{2} \rceil$, for all k.

For the last class, φ is accepted if the difference between those that agree and those that disagree with it exceed a threshold that depends on the number of abstentions.

3.21. Definition. *Consider a function $vm : \{0, \dots, n\} \times \Phi \to \{-n, \dots, n+1\}$. Denote the number $vm(k, \varphi)$ by m^k_φ. The* ***variable marginal quota rule*** *F_{vm} is such that:*

$$F_{vm}(\boldsymbol{J}) = \{\varphi \in \Phi \mid n^{\boldsymbol{J}}_\varphi - n^{\boldsymbol{J}}_{\neg\varphi} \geqslant m^{n^{\boldsymbol{J}}_{\widetilde{\varphi}}}_\varphi\}$$

In the European Union Council, an abstention on a matter decided by unanimity (e.g., a matter on taxation, family law, or citizenship) has the effect of a *yes* vote:[13] A decision comes into force if and only if all non-abstaining votes are in favour of it. Concretely, we have a variable marginal threshold $m^k_\varphi = k$, for all k.

Other commonly used rules can also be described within the model of quota rules presented above. For example:

- The absolute majority uses an invariable absolute threshold $a_\varphi = \lceil \frac{n}{2} \rceil$.
- A procedure with a quorum r (where a decision goes through if and only if at least r agents do not abstain and from those a majority supports the decision under discussion) is equivalent to the implementation of a variable absolute threshold such that $a^k_\varphi = n + 1$ if $k < r$ and $a^k_\varphi = \lceil \frac{k}{2} \rceil$ if $k \geqslant r$.

For the variable absolute (or marginal) quota rules and for any specific threshold a^k_φ (or m^k_φ), if the threshold is at least $k + 1$, then proposition φ will never be accepted when k judgments about the issue $\widetilde{\varphi}$ are reported. This means that in practice, any rule with threshold larger than k will behave the same. Analogously, for variable marginal quota rules, all thresholds smaller than $-k$ will effectively be the same, meaning that the relevant proposition will always be accepted when k judgments on $\widetilde{\varphi}$ are reported.

> The definitions we use may seem to lack in elegance, including multiple thresholds with the same function, but are indispensable for establishing that the variable quota rules constitute the most general class of rules.

Moreover, the above definitions include a class of (invariable) rules called *trivial*: Trivial rules are such that, for all propositions $\varphi \in \Phi$, one of the following holds for all numbers $k = n^{\boldsymbol{J}}_{\widetilde{\varphi}}$ of reported judgments: (i) φ is always accepted (that is, the relevant absolute (marginal) threshold is equal to 0 ($-n$)), (ii) φ is always rejected (that is, the

[12] https://web.archive.org/web/20170110092314/https://www.ch.ch/en/referendum

[13] https://web.archive.org/web/20200429153617/https://www.consilium.europa.eu/en/council-eu/voting-system/unanimity/

relevant (absolute or marginal) threshold is equal to $n + 1$), (*iii*) φ is accepted if and only if there is a unanimous support in favour of it (that is, the relevant (absolute or marginal) threshold is equal to n).

Note also that we only consider quota rules where the same type of threshold is associated with every proposition—in particular, we cannot have an invariable absolute threshold for some proposition φ and an invariable marginal threshold for a different proposition ψ within the same rule.

Example 3.22 illuminates the difference between the weight rules for judgment aggregation that we defined in Section 3.1.1, and the quota rules of this section.

3.22. Example. Let us come back to Example 3.17 on page 34 and suppose that we have ten agents. Take on the one hand the equal-and-even weight rule, and on the other hand a marginal quota rule prescribing that a positive proposition be collectively accepted if and only if at least two agents accept it (that is, 20% of the group), while a negative proposition needs the support of at least nine agents to support it (that is, at least 85% of the group). The equal-and-even weight rule will say *no* to farm B (because the relevant proposition gets weight 1 from eight agents, which is more than the weight of any other combination of propositions that the rule can choose to accept in order to agree with the two remaining agents), and it will also say *no* to fruit on campus. On the contrary, the described quota rule will follow the will of the two agents, agreeing with the distribution of fruit on campus, and selecting farm B as the supplier. $\triangle$

Relations between Different Quota Rules. Besides enlisting a number of different quota rules for incomplete judgments, it is also interesting to explore the connections between these rules. Intriguing questions that we shall answer read as follows:

> Do all definitions we have so far provided generate distinct rules? Is one particular way of describing quota rules in judgment aggregation the most general one? If some definitions induce equivalent rules, to what classes do these rules belong?

Figure 3.1 responds to our questions graphically, depicting the space of quota rules for incomplete judgments on which we will shortly elaborate.

We next prove all relevant relations between our classes of rules. To start, variable quota rules clearly are more general than invariable quota rules, since variable thresholds can simply be constant in order to capture invariable ones.

But what about the relations between absolute and marginal thresholds? As we will see, the answers here differ completely across the classes of variable and invariable quota rules: invariable ones, divided into absolute and marginal, are proven to define two disjoint classes of rules (except for trivial cases), while variable ones, absolute and marginal, define exactly the same class of rules.

3.23. Proposition. *There is no non-trivial invariable absolute quota rule that coincides with an invariable marginal quota rule (and vice versa).*

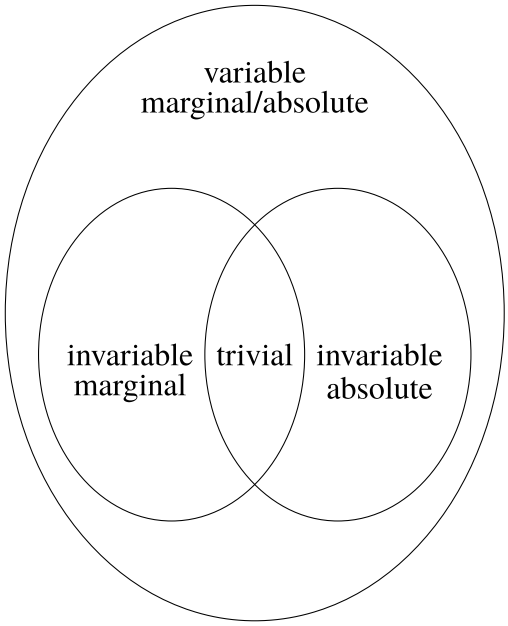


Figure 3.1: The space of quota rules for incomplete judgments.

Proof. Take an arbitrary non-trivial invariable absolute quota rule F_a, induced by the quota function a. For some $\varphi \in \Phi$, it is the case that $a_\varphi \neq 0$, $a_\varphi \neq n+1$, and $a_\varphi \neq n$ (thus, $1 \leqslant a_\varphi \leqslant n-1$). Suppose—aiming for a contradiction—that F_a coincides with some invariable marginal quota rule F_m, induced by the quota function m. Consider a profile $\boldsymbol{J}$ where exactly a_φ agents accept φ and one agent accepts $\neg\varphi$ (i.e., $n^{\boldsymbol{J}}_\varphi = a_\varphi$ and $n^{\boldsymbol{J}}_{\neg\varphi} = 1$)—this is possible because $a_\varphi + 1 \leqslant n$. Then, it holds that $\varphi \in F_a(\boldsymbol{J})$, so it must also be the case that $\varphi \in F_m(\boldsymbol{J})$. But for $\varphi \in F_m(\boldsymbol{J})$ to hold, we must have that $m_\varphi \leqslant n^{\boldsymbol{J}}_\varphi - n^{\boldsymbol{J}}_{\neg\varphi} = a_\varphi - 1$.

Now, since $a_\varphi - 1 \geqslant 0$, we consider a different profile $\boldsymbol{J}'$ where exactly $a_\varphi - 1$ agents accept φ and no-one accepts $\neg\varphi$: $n^{\boldsymbol{J}'}_\varphi = a_\varphi - 1$ and $n^{\boldsymbol{J}'}_{\neg\varphi} = 0$. Since $m_\varphi \leqslant n^{\boldsymbol{J}'}_\varphi - n^{\boldsymbol{J}'}_{\neg\varphi} = a_\varphi - 1$, we have that $\varphi \in F_m(\boldsymbol{J}')$. Thus, we must also have that $\varphi \in F_a(\boldsymbol{J}')$. But $n^{\boldsymbol{J}'}_\varphi = a_\varphi - 1 < a_\varphi$, which means that $\varphi \notin F_a(\boldsymbol{J}')$, implying a contradiction. □

3.24. Proposition. *Every variable absolute quota rule coincides with some variable marginal quota rule, and every variable marginal quota rule coincides with some variable absolute quota rule.*

Proof. Consider a variable absolute quota rule F_{va}, associated with thresholds a^k_φ. We need to construct a suitable variable marginal quota rule F_{vm} (associated with thresholds m^k_φ) and show that $F_{va}(\boldsymbol{J}) = F_{vm}(\boldsymbol{J})$ for all profiles $\boldsymbol{J}$. For all propositions φ and numbers k, we define the following thresholds:

$$m^k_\varphi = \begin{cases} k+1 & \text{if } a^k_\varphi \geqslant k+1 \\ 2a^k_\varphi - k & \text{otherwise} \end{cases}$$

Take an arbitrary profile $\boldsymbol{J}$ and a proposition $\varphi \in \Phi$ with $\varphi \in F_{va}(\boldsymbol{J})$. By definition of the variable absolute quota rule, $\varphi \in F_{va}(\boldsymbol{J})$ means that $n^{\boldsymbol{J}}_{\varphi} \geqslant a_{\varphi}^{n^{\boldsymbol{J}}_{\widetilde{\varphi}}}$. But we also have that $n^{\boldsymbol{J}}_{\varphi} + n^{\boldsymbol{J}}_{\neg\varphi} = n^{\boldsymbol{J}}_{\widetilde{\varphi}}$. Thus, $n^{\boldsymbol{J}}_{\varphi} \geqslant a_{\varphi}^{n^{\boldsymbol{J}}_{\widetilde{\varphi}}}$ implies that $n^{\boldsymbol{J}}_{\neg\varphi} \leqslant n^{\boldsymbol{J}}_{\widetilde{\varphi}} - a_{\varphi}^{n^{\boldsymbol{J}}_{\widetilde{\varphi}}}$, from which we can derive that $n^{\boldsymbol{J}}_{\varphi} - n^{\boldsymbol{J}}_{\neg\varphi} \geqslant 2a_{\varphi}^{n^{\boldsymbol{J}}_{\widetilde{\varphi}}} - n^{\boldsymbol{J}}_{\widetilde{\varphi}} = m_{\varphi}^{n^{\boldsymbol{J}}_{\widetilde{\varphi}}}$ (since φ is accepted by F_{va} in the profile $\boldsymbol{J}$, we know that $a^k_{\varphi}, m^k_{\varphi} \leqslant k$). Hence, by definition of the variable marginal quota rule we conclude that $\varphi \in F_{vm}(\boldsymbol{J})$ and so that $F_{va}(\boldsymbol{J}) \subseteq F_{vm}(\boldsymbol{J})$.

Now, take a proposition $\varphi \in \Phi$ with $\varphi \in F_{vm}(\boldsymbol{J})$. By definition of the variable marginal quota rule, $\varphi \in F_{vm}(\boldsymbol{J})$ means that $n^{\boldsymbol{J}}_{\varphi} - n^{\boldsymbol{J}}_{\neg\varphi} \geqslant m_{\varphi}^{n^{\boldsymbol{J}}_{\widetilde{\varphi}}}$. Since we also know that $n^{\boldsymbol{J}}_{\varphi} + n^{\boldsymbol{J}}_{\neg\varphi} = n^{\boldsymbol{J}}_{\widetilde{\varphi}}$, we obtain that $2n^{\boldsymbol{J}}_{\varphi} \geqslant m_{\varphi}^{n^{\boldsymbol{J}}_{\widetilde{\varphi}}} + n^{\boldsymbol{J}}_{\widetilde{\varphi}} = 2a_{\varphi}^{n^{\boldsymbol{J}}_{\widetilde{\varphi}}} - n^{\boldsymbol{J}}_{\widetilde{\varphi}} + n^{\boldsymbol{J}}_{\widetilde{\varphi}}$ (again, since φ is accepted by F_{vm} in the profile $\boldsymbol{J}$, we know that $m^k_{\varphi}, a^k_{\varphi} \leqslant k$). Equivalently, $n^{\boldsymbol{J}}_{\varphi} \geqslant a_{\varphi}^{n^{\boldsymbol{J}}_{\widetilde{\varphi}}}$, which implies that $\varphi \in F_{va}(\boldsymbol{J})$. Thus, we also have that $F_{(m^k_{\varphi})}(\boldsymbol{J}) \subseteq F_{va}(\boldsymbol{J})$, concluding that $F_{va}(\boldsymbol{J}) = F_{vm}(\boldsymbol{J})$.

The fact that every variable marginal quota rule coincides with some variable absolute quota rule can be proven in a directly analogous manner. □

Thus, when considering quota rules for the aggregation of incomplete individual judgments, it suffices to study variable absolute quota rules—all obtained results will also hold for variable marginal quota rules.

Collective Rationality Properties of Quota Rules. Three[14] properties of aggregation rules F that capture different notions of *collective rationality* have been considered in the literature, talking about desirable forms of the collective outcomes.[15]

Completeness requires that $\varphi \in F(\boldsymbol{J})$ or $\neg\varphi \in F(\boldsymbol{J})$, for all profiles $\boldsymbol{J}$ and propositions $\varphi \in \Phi$. *Complement-freeness* requires that $\varphi \in F(\boldsymbol{J})$ and $\neg\varphi \in F(\boldsymbol{J})$, for all profiles $\boldsymbol{J}$ and propositions $\varphi \in \Phi$. Lastly, *consistency* requires that $F(\boldsymbol{J})$ be logically consistent, for all profiles $\boldsymbol{J}$.

3.25. Example. Consider a profile regarding two issues $\widetilde{p}$ and $\widetilde{q}$, as follows:

$$J_1 = \{p, q\}$$
$$J_2 = \{\neg p, q\}$$
$$J_3 = \{\neg q\}$$

[14] A fourth rationality property that we do not discuss here requires that the judgment sets be *deductively closed*, i.e., that they contain all propositions that are logical consequences of their members.

[15] We use the term "collective rationality" referring to properties that are compelling for the outcome of an aggregation process, but these properties are not necessarily desirable for individual judgments. In particular, we may be interested in outcomes that are *complete*, i.e., that make a decision on each issue at stake, although individual inputs may still be incomplete.

Suppose that we employ the absolute majority rule for both issues, which collectively accepts a proposition if and only if more than half of the agents accept it. Then, the collective outcome would be $\{q\}$—since no collective decision is made about the issue $\widetilde{p}$, the rule is not complete. To encourage completeness, suppose that we select a rule that accepts a proposition if at least one agent includes it in her judgment. Then, the collective outcome would be $\{p, \neg p, q\}$. But since both a negative and a positive opinion about $\widetilde{p}$ appear in the collective decision, this new rule is not complement-free. Obviously, every rule that is not complement-free is not consistent either. △

As we already know, notions of collective rationality such as completeness, complement-freeness, and consistency are not necessarily satisfied by quota rules.[16]

In the complete case, a prominent example is presented by the majority rule, which collectively accepts a proposition when more than half of the agents agree with this decision. The majority rule is complete for an odd number of agents and complete inputs. However, it generates logical inconsistencies even in slightly more complex agendas, like the one instantiated in the famous *discursive dilemma* illustrated in Table 2.1, in Chapter 2, page 12. Thus, a natural question arises:

> Can problematic cases of inconsistent collective outcomes be avoided by allowing agents to report incomplete judgments?

Although this direction may look promising at first, we will soon see that individual incompleteness is in fact detrimental to logical consistency, as soon as we require complete outcomes (indeed, if completeness of the collective decision is dropped, then inconsistencies can be trivially resolved).

3.26. Example. Consider the simple majority rule, which accepts a proposition φ if more agents report φ than $\neg\varphi$. This rule is guaranteed to induce complete outcomes only in the special cases where an odd number of agents submit a judgment on $\widetilde{\varphi}$. Still, the discursive dilemma persists (consult Table 3.2). Note that this example does not go through for the absolute majority rule. △

Even worse, inconsistencies appear under incompleteness also for very simple agendas where only two propositions are incompatible together.

3.27. Example. Consider an agenda $\Phi = \{\varphi, \psi, \neg\varphi, \neg\psi\}$, with $\{\varphi, \psi\}$ being inconsistent, and see Table 3.3 for an inconsistent collective decision. This example works specifically for the simple (not the absolute) majority rule. △

3.28. Proposition. *For any agenda Φ that does not only consist of logically independent issues, the simple majority rule for incomplete individual judgments is inconsistent.*

[16] An important difference between weight rules and quota rules is that the former guarantee complete and consistent collective judgments, while the latter do not. However, in the special case of logically independent propositions and of complete individual judgments, weight rules reduce to a very familiar quota rule: the majority rule.

	p	q	$p \wedge q$
Agent 1:	☑	☐	☐
Agent 2:	☒	☐	☒
Agent 3:	☐	☑	☐
Agent 4:	☑	☐	☐
Simple majority:	☑	☑	☒

Table 3.2: The discursive dilemma with incomplete individual judgments.

	φ	ψ
Agent 1:	☑	☐
Agent 2:	☒	☑
Agent 3:	☑	☐
Simple majority:	☑	☑

Table 3.3: A logically inconsistent outcome $\{\varphi, \psi\}$ for incomplete individual judgments on a simple agenda.

Proof. Consider an agenda Φ as in the statement, and an inconsistent set $Z \subseteq \Phi$ such that every strict subset of Z is consistent (such sets are called *minimally inconsistent*). Take two propositions $\varphi, \psi \in Z$ such that $\psi \neq \neg\varphi$. Construct the profile where an agent i accepts all propositions in Z besides φ, abstaining on everything else, and an agent j accepts all propositions in Φ besides ψ, abstaining on everything else. The remaining agents abstain on all issues. Then, the simple majority rule will include all propositions of Z in the collective outcome, making it inconsistent. □

Yet, some variable quota rules present immediate advantages in contexts where consistency, together with completeness, is desired. Consider for instance the simple agenda of Table 3.3, and assume that the group of agents is of an odd number. Then, there exists a variable quota rule that is complete and consistent, and functions as follows: it always accepts φ and $\neg\psi$ as the collective decision, besides the case where all agents report their judgments on both issues; in the latter case of unanimous participation, the rule follows the opinion of the majority on each issue.

A more detailed analysis of the conditions on the relevant thresholds that guarantee properties of collective rationality for the associated quota rules has been conducted by Terzopoulou (2020). The obtained results also imply that verifying whether a quota rule is collectively rational under a given interpretation is never harder in the more general incomplete setting compared to the framework of complete inputs. Thus, loosely speaking, the framework's expressivity can be increased without paying a price in terms of increased complexity.

3.3 Positional Scoring Rules in Voting

In this section we will get acquainted with the class of positional scoring rules for voting with incomplete preferences. Positional scoring rules, which hinge on giving points to alternatives depending on the position they occupy in an agent's preference, have been defined and widely used in the traditional setting with complete preferences—they are very easy to understand, and straightforward to compute. In our broader framework, these characteristics still apply. However, a more flexible notion of "position" is needed when we talk about preferences that take the form of general directed acyclic graphs rather than plain rankings. We give our definitions for arbitrary preferences $R \in \mathcal{D}(A)$.

A *scoring function* is a function $s\colon (\mathcal{D}(A) \times A) \to \mathbb{R}$ that assigns a score to every alternative in a given preference, for all sets of alternatives $A \subseteq A^\star$ and groups $N \subseteq N^\star$. We abbreviate $s(R, a)$ by $s_R(a)$. A scoring function is *non-trivial* if $s_R(x) \neq s_R(y)$ for some $x, y \in A$ and $R \in \mathcal{D}(A)$. Given a profile $\boldsymbol{R} = (R_1, \ldots, R_n)$, we will often also use the abbreviation $s_{\boldsymbol{R}}(a) = \sum_{i \in N} s_{R_i}(a)$.

A *positional scoring function* ensures the symmetrical treatment of all alternatives; one may thus think of scores assigned to positions in a graph (see Figure 3.2 for an example). Formally, a scoring function s is positional if and only if for all permutations $\sigma\colon A \to A$, preferences $R \in \mathcal{D}(A)$, and alternatives $x \in A$, we have that $s_R(x) = s_{R_\sigma}(x_\sigma)$, where $x_\sigma = \sigma(x)$ and $R_\sigma = \sigma(R) = \{(a_\sigma, b_\sigma) \mid aRb\}$.

$$\begin{tikzcd} a \arrow[d] \arrow[dr] & c \\ b & d \end{tikzcd} \qquad \begin{tikzcd} 3 \arrow[d] \arrow[dr] & 0 \\ 2 & 2 \end{tikzcd} \qquad \begin{tikzcd} 3 \arrow[d] \arrow[dr] & 0 \\ 2 & 1 \end{tikzcd}$$

Figure 3.2: The center and right graphs represent possible scoring functions of the graph on the left. The scoring function represented by the center graph is positional. The one on the right is not positional, because the two alternatives b and d are assigned different scores, while appearing in positions that are indistinguishable under permutation.

A scoring rule for incomplete preferences F_s, associated with a scoring function s, takes a profile of preferences for some $A \subseteq A^\star$ and $N \subseteq N^\star$, and returns all alternatives with maximal scores, where the score of an alternative is the sum of the scores across all agents. Formally, $F_s\colon \mathcal{D}(A)^n \to 2^A \setminus \{\emptyset\}$ is defined as follows:

$$F_s(\boldsymbol{R}) \quad = \quad \operatorname*{argmax}_{x \in A} \sum_{i \in N} s_{R_i}(x) \quad = \quad \operatorname*{argmax}_{x \in A} s_{\boldsymbol{R}}(x)$$

Note that for distinct scoring functions s and s' it may be the case that $F_s = F_{s'}$. Then, we say that s and s' are *equivalent*. More specifically, given an infinite superpopulation, two scoring functions s and s' are equivalent if and only if there are real numbers $\{\alpha_R\}_{R\in\mathcal{P}(A)}$ and $\beta > 0$ such that for all alternatives $x \in A$ and preferences R, $s_R(x) = \alpha_R + \beta \cdot s'_R(x)$.

> An infinite superpopulation is essential in order to guarantee that a scoring function that is not a positive affine transformation of another one, but is close to it, will give rise to a different collective outcome for some preference profile.

3.3.1 A Few Selected Rules

Let us now see some concrete positional scoring functions, specifically defined for strict incomplete preferences $P \in \mathcal{P}(A)$. We start off with two classes of such functions—the *approval* class and the *veto* class—that generalise classical approval rules from standard voting with linear orders. Given $P \in \mathcal{P}(A)$ and $x \in A$, for $k \in \{1, \ldots, m\}$, the k-approval rule assigns to x score 0 if there exist at least k alternatives above x, and score 1 otherwise; the k-veto rule assigns to x score 1 if there exist at least k alternatives below x, and score 0 otherwise.

For complete preferences, k-approval (veto) entails that *precisely k* alternatives are approved (disapproved), while for incomplete preferences, k-approval (veto) implies that *at least k* alternatives are approved (disapproved). Furthermore, for complete profiles k-approval and $(m-k)$-veto coincide (but this is not true for incomplete ones—see Figure 3.3 for an example); also, 1-approval and $(m-1)$-veto correspond to the plurality rule, while 1-veto and $(m-1)$-approval reduce to the antiplurality rule.

$$
\begin{array}{cc}
a & c \\
\downarrow & \downarrow \\
b & d
\end{array}
\qquad\qquad
\begin{array}{cc}
1 & 1 \\
\downarrow & \downarrow \\
0 & 0
\end{array}
\qquad\qquad
\begin{array}{cc}
0 & 0 \\
\downarrow & \downarrow \\
0 & 0
\end{array}
$$

Figure 3.3: The scores assigned by the 1-approval scoring function (middle) and by the 3-veto scoring function (right), given a strict preference (left) over four alternatives. The 1-approval rule selects as winners the two alternatives on top, a and c, but the 3-veto rule is indecisive between all four alternatives.

3.29. DEFINITION. *Given some number $k \in \{1, \ldots, m\}$, define the **k-approval** scoring function $s^{k\text{-approval}}\colon (\mathcal{P}(A) \times A) \to \mathbb{R}$ as follows:*

$$s_P^{k\text{-approval}}(x) = \begin{cases} 1 & \text{if } |\{y \in A \mid yPx\}| < k \\ 0 & \text{if } |\{y \in A \mid yPx\}| \geqslant k \end{cases}$$

3.30. DEFINITION. *Given some number $k \in \{1, \ldots, m\}$, define the **k-veto** scoring function $s^{k\text{-veto}}\colon (\mathcal{P}(A) \times A) \to \mathbb{R}$ as follows:*

$$s_P^{k\text{-veto}}(x) = \begin{cases} 1 & \text{if } |\{y \in A \mid xPy\}| \geqslant k \\ 0 & \text{if } |\{y \in A \mid xPy\}| < k \end{cases}$$

We continue with the definition of a few more positional scoring functions for incomplete preferences, which do not belong in the two classes just presented.

3.31. DEFINITION. *The **domination scoring** function* $\mathrm{ds}\colon (\mathcal{P}(A) \times A) \to \mathbb{R}$ *is defined as follows:*

$$\mathrm{ds}_P(x) = |\{y \in A \mid xPy\}|$$

3.32. DEFINITION. *The **cumulative scoring** function* $\mathrm{cs}\colon (\mathcal{P}(A) \times A) \to \mathbb{R}$ *is defined recursively, as follows:*

$$\mathrm{cs}_P(x) = \begin{cases} 0 & \textit{if } xPy \textit{ for no } y \in A \\ 1 + \sum_{\substack{y \in A \\ xPy}} \mathrm{cs}_P(y) & \textit{otherwise} \end{cases}$$

For the next scoring function, a new definition is first in order. Given a preference P, we say that two alternatives a and b are *connected* if $a = b$, or if there is a path from a to b in the undirected version of the graph defined by P. The alternatives a and b are said to be in an *undirected cycle* if that graph has a cycle that contains both, or if $a = b$.

3.33. DEFINITION. *The **stepwise scoring** function* $\mathrm{ss}\colon (\mathcal{P}(A) \times A) \to \mathbb{R}$ *is defined by*

$$\mathrm{ss}_P(x) = \sum_{i=1}^{k} \overline{s}(C_i, C),$$

where $\mathcal{C} = \{C_1, \ldots, C_k\}$ *is a partition of* A *into sets of alternatives that are in the same undirected cycle,* $x \in C$ *for* $C \in \mathcal{C}$*, and* $\overline{s} : (\mathcal{C}\times\mathcal{C}) \to \mathbb{R}$ *is defined below (for* P^c_u *denoting the undirected version of the relation* $P^c \subseteq \mathcal{C}\times\mathcal{C}$ *such that* $C_i P^c C_j$ *if and only if* xPy *for some* $x \in C_i$ *and* $y \in C_j$*):*[17]

If $C_i = C_j$*, then* $\overline{s}(C_i, C_j) = 0$*, and if* $C_i \neq C_j$*, then we have the following:*

$$\overline{s}(C_i, C_j) = \begin{cases} 1/2 & \textit{if } C_i P^c_u C_1 P^c_u \ldots C_m P^c C_j \textit{ for some } m \geqslant 0 \\ -1/2 & \textit{if } C_j P^c C_1 P^c_u \ldots C_m P^c_u C_i \textit{ for some } m \geqslant 0 \\ 0 & \textit{otherwise} \end{cases}$$

Stepwise scoring assigns to two alternatives x and y the same score when they are in the same undirected cycle. Otherwise, if x is preferred to y, then $\mathrm{ss}_P(x) = \mathrm{ss}_P(y) + 1$.[18] The cumulative scoring function gives each alternative a score one greater than the summed score of all the alternatives to which it is preferred. For complete input profiles, the domination scoring rule reduces to the Borda rule; the cumulative scoring rule reduces to the (complete) positional scoring rulewith score vector $(2^{|A|-2}, \ldots, 4, 2, 1, 0)$; the veto scoring rule reduces to the veto rule; and the stepwise scoring rule reduces to the trivial rule that returns the set of all alternatives for all input profiles. See Figure 3.4 for an illustration of the different scores these functions give for an example preference.

[17]Note that the stepwise scores of the alternatives in a preference P may need to be normalised in order to ensure that $\min\{s_P(x) : x \in A\} = 0$.

[18]Note that $\overline{s}$ is well-defined, because the graph of $\overline{s}$ over $\mathcal{C}\times\mathcal{C}$ is a forest, and hence the three cases in the relevant definition are the only possible ones.

$$
\begin{tikzcd}
1 \arrow[r] & 1 \arrow[r] & 0 \\
3 \arrow[u] \arrow[r] \arrow[d] & 2 \arrow[r] \arrow[ul] & 1 \arrow[d] \\
0 & 1 \arrow[ur] & 0
\end{tikzcd}
\qquad
\begin{tikzcd}
2 \arrow[r] & 1 \arrow[r] & 0 \\
7 \arrow[u] \arrow[r] \arrow[d] & 4 \arrow[r] \arrow[ul] & 1 \arrow[d] \\
0 & 2 \arrow[ur] & 0
\end{tikzcd}
\qquad
\begin{tikzcd}
2 \arrow[r] & 1 \arrow[r] & 0 \\
2 \arrow[u] \arrow[r] \arrow[d] & 2 \arrow[r] \arrow[ul] & 1 \arrow[d] \\
1 & 2 \arrow[ur] & 0
\end{tikzcd}
$$

Figure 3.4: Left to right; scores given by domination, cumulative, and stepwise scoring for an example preference.

An intriguing link exists between the class of positional scoring rules we have introduced in this section and the weight rules that produce single winners (recall Section 3.1.2). Specifically, the latter constitute special cases of the former. The basic intuition can be gained by inspecting the special case of complete individual preferences, where single-winner weight rules reduce to the Borda rule: a renowned positional scoring rule (Proposition 3.14). In the general case of incomplete preferences, single-winner weight rules can be regarded as a weighted variant of the domination scoring rule (Definition 3.31). We explain our reasoning formally:

Call $P^x \in \mathcal{W}(A)$ the preference that identifies the alternative $x \in A$ as the winner, and for a set of such preferences $S \subseteq \mathcal{W}(A)$, denote by *top-alt*$(S) \subseteq A$ all those alternatives that are identified as the winner by some preference in S. Then, the winning alternatives according to a weight rule of single-winner type (associated with a weight function w) are the following:

$$
\begin{aligned}
\textit{top-alt}\big(F_w^{\mathcal{W}}(\boldsymbol{P})\big) \quad &= \quad \textit{top-alt}\big(\operatorname*{argmax}_{P^x \in \mathcal{W}(A)} \sum_{i \in N} w_{|P_i|} \cdot |P^x \cap P_i| \big) \\
&= \quad \textit{top-alt}\big(\operatorname*{argmax}_{P^x \in \mathcal{W}(A)} \sum_{i \in N} w_{|P_i|} \cdot |y \in A \mid xP_i y| \big) \\
&= \quad \operatorname*{argmax}_{x \in A} \sum_{i \in N} w_{|P_i|} \cdot |y \in A \mid xP_i y| \\
&= \quad \operatorname*{argmax}_{x \in A} \sum_{i \in N} w_{|P_i|} \cdot \mathrm{ds}_{P_i}(x)
\end{aligned}
$$

As an illustration, see Figure 3.5, where the equal-and-even weight rule is applied. The scores corresponding to each alternative are depicted next to that alternative.

$$
\begin{tikzcd}
1/2+1/2 & & 0 \\
a \arrow[rr] \arrow[dr] & & c \\
& b & \\
& 0 &
\end{tikzcd}
\qquad
\begin{tikzcd}
1/3 & & 0 \\
a \arrow[rr] & & c \\
& b \arrow[ul] \arrow[ur] & \\
& 1/3 + 1/3 &
\end{tikzcd}
\qquad
\begin{tikzcd}
0 & & 1 \\
a & & c \arrow[dl] \\
& b & \\
& 0 &
\end{tikzcd}
$$

Figure 3.5: Example of equal-and-even scores in a preference profile.

3.3.2 The Borda Class

Let us now exclusively focus on the Borda rule. This is a very well-established positional scoring rule, applied to profiles of complete preferences over m alternatives. Recall that it prescribes for each voter that $m-1$ points be given to her top alternative, $m-2$ points to her second-to-top alternative, and so forth, with 0 points given to the alternative ranked last. The alternatives with the largest sum of points across all voters are then announced the winners of the election.

What are appropriate extensions of the Borda rule for incompleteness?

We address this question for domains of preferences that are *top-truncated*. A preference is called top-truncated if it consists of a linear order over a subset of the given alternatives, with the assumption that all non-ranked alternatives are inferior to all ranked alternatives. Top-truncated preferences provide a sensible model for real-world applications. From choosing the members of a parliament to selecting favourite movies to add to a watch-list, an agent is likely to be more interested in her most preferred alternatives, and thus also more sensitive to their similarities and differences, deducing a full ranking among these alternatives.

In this thesis, we specifically consider two domains of top-truncated preferences, $\mathcal{D}_1(A)$ and $\mathcal{D}_2(A)$. In each of our domains, all preferences take the same form, which is one of those depicted in Figure 3.6 (where transitive arrows are omitted for simplicity).

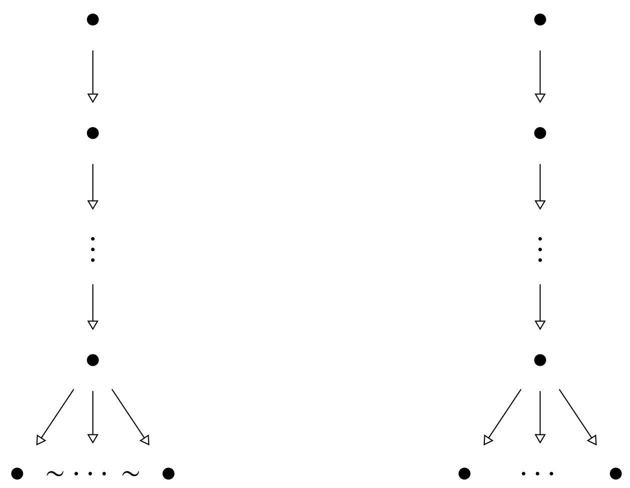


Figure 3.6: Types of top-truncated preferences in two domains: $\mathcal{D}_1(A)$ on the left and $\mathcal{D}_2(A)$ on the right.

In particular, $\mathcal{D}_1(A)$ contains top-truncated preferences where k alternatives are ranked, for any $0 \leqslant k \leqslant m$, and the alternatives that are not ranked are indistinguishable from each other; $\mathcal{D}_2(A)$ also contains top-truncated preferences where k alternatives are ranked, for any $0 \leqslant k \leqslant m$, but the alternatives that are not ranked are incomparable to each other. Note that $\mathcal{D}_1(A) \subset \mathcal{D}(A)^t$ and $\mathcal{D}_2(A) \subset \mathcal{P}(A)^t$. Formally, a top-truncated preference $\succsim$ consists of two parts: the *top* the *bottom*. Let us define $TOPsets(\succsim)$ as the collection of all subsets $A' \subseteq A$ that contain strictly ordered alternatives in $\succsim$ that are superior, according to $\succsim$, to all alternatives not in A':

$$TOPsets(\succsim) = \{A' \subseteq A \mid (i)\ x \succ z \text{ for all } x \in A', z \in A \setminus A' \text{ and } \\ (ii)\ x \succ y \text{ or } y \succ x \text{ for all } x, y \in A'\}$$

$TOP(\succsim)$ is the unique largest set in $TOPsets(\succsim)$, and $BOT(\succsim) = A \setminus TOP(\succsim)$. The set of top alternatives in a profile $\boldsymbol{\succsim}$ includes those alternatives that are on top for all agents. That is, $TOP(\boldsymbol{\succsim})$ is defined as the unique largest set in $\bigcap_{i \in N} TOPsets(\succsim_i)$ if that intersection is non-empty; otherwise, $TOP(\boldsymbol{\succsim})$ is the empty set. We also define $BOT(\boldsymbol{\succsim})$ as the set $A \setminus TOP(\boldsymbol{\succsim})$, but note that this defnition is intuitively meaningful (in the sense of characterising a set of "bottom alternatives") only when $TOP(\boldsymbol{\succsim}) \neq \emptyset$.

See Figure 3.7 for an example.

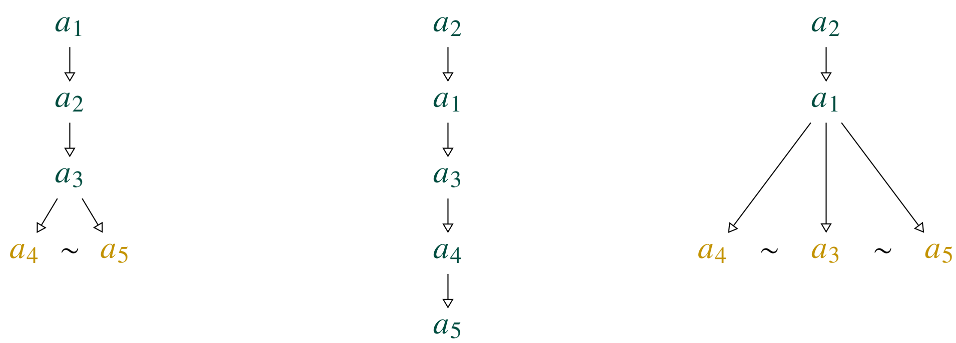


Figure 3.7: The top alternatives of each single preference are in blue. From left to right, these are the sets $\{a_1, a_2, a_3\}$, $\{a_1, \ldots, a_5\}$, and $\{a_1, a_2\}$, respectively. The $TOPsets(\succsim)$ of the leftmost preference $\succsim$ are $\{\{a_1\}, \{a_1, a_2\}, \{a_1, a_2, a_3\}\}$, and similarly for the other preferences. The intersection of these sets is then the set $\{a_1, a_2\}$, which includes the top alternatives of the whole profile.

The Borda rule is commonly defined on domains of linear orders in one of two ways. First, as a positional scoring rule with score-vector $(m-1, m-2, \ldots, 0)$; second, in terms of the *weighted majority graph*, where the winning alternatives are those that maximise the following function:[19]

$$B(a) \quad = \quad \sum_{y \in A} |\{i \in N \mid a \succ_i y\} - |\{i \in N \mid y \succ_i a\}|| \quad = \quad \sum_{i \in N} \sum_{y \in A} \mathbb{1}_{a \succ_i y} - \mathbb{1}_{y \succ_i a}$$

We can think of $B(a)$, the *symmetric Borda score* of a, as $B(a) = \sum_{i \in N} B_i(a)$ with $B_i(a) = \sum_{y \in A} \mathbb{1}_{a \succ_i y} - \mathbb{1}_{y \succ_i a}$.

[19]The definition of the Borda rule based on the weighted majority graph can also be considered a definition of a positional scoring rule on linear orders, with score-vector $(m, m-2, \ldots, -m)$.

It seems sensible to presuppose that any interesting generalisation of the Borda rule will also be defined in terms of a scoring function B with $B(x) = \sum_{i \in N} B_i(x)$ for some scoring functions B_i that each only makes reference to the preference of one agent i. In Figure 3.8 we present some options for how one could define such a function on general domains of incomplete preferences allowing for indifference—all of them reduce to the standard Borda rule for profiles of linear orders.

$B_i(a) = \sum_{y \in A} \mathbb{1}_{a \succ_i y} - \mathbb{1}_{y \succ_i a}$

(a) the symmetric Borda score.

$B_i(a) = \sum_{y \in A} \mathbb{1}_{a \succ_i y}$

(b) based on strict dominance.

$B_i(a) = \sum_{y \in A} \mathbb{1}_{a \succsim_i y}$

(c) based on weak dominance.

$B_i(a) = -\sum_{y \in A} \mathbb{1}_{y \succ_i a}$

(d) based on non strict dominance.

Figure 3.8: Domination-based scores for defining the Borda rule on (possibly) incomplete preferences. When indistinguishability is not materialised in a domain, definition (b) coincides with definition (c); when incomparability is not materialised, definition (c) coincides with definition (d).

Some suggestions regarding appropriate generalisations of the Borda rule for top-truncated preferences have been made in the literature already. All of them are reasonable at first sight, but heavily depend on the interpretation of the preference domain, and of the rule, we have in mind. For instance, given a preference with k ranked alternatives out of a total of m alternatives overall, should the points assigned to the unranked alternatives at the bottom be $m - k - 1$ (as if they all were ranked at level $k + 1$ from the top), or should the points be 0 (as if all unranked alternatives were ranked at the very lowest level m)? Dummett (1997), Saari (2008), Baumeister et al. (2012), Cullinan et al. (2014), and Caragiannis et al. (2015) have presented their own versions of the Borda rule (some of which coincide). Emerson (2013) has informally discussed the advantages and disadvantages of different such generalisations of the Borda rule, concentrating on issues related to strategic behaviour.

All generalisations of the Borda rule that have appeared in previous work follow the definition of a positional scoring rule in terms of score-vectors. Note that for domains of top-truncated preferences we actually need $m - 1$ such vectors, one for each possible number of the top alternatives in a preference (when this number is 0, all alternatives will always be assigned the same score, by the definition of positional scoring functions). Now, the last $m - k$ positions in a vector correspond to the scores associated with the bottom alternatives in a top-truncated preference with k top alternatives. Note that all bottom alternatives must have the same score, by the definition of positional scoring functions. Generally, given a top-truncated preference $\succsim$ with $|TOP(\succsim)| = k \leqslant m - 1$, we will write s_j for the score of the alternative ranked in the jth position within the top part of $\succsim$ and s_{k+1} for the score of all alternatives in the bottom part of $\succsim$. These are the versions of the Borda rule for top-truncated preferences in the literature:

- **Pessimistic Borda** (Baumeister et al., 2012):[20]

$$(m-1, m-2, \ldots, m-k, 0, \ldots, 0), \text{ for all } 1 \leqslant k \leqslant m$$

- **Optimistic Borda** (Saari, 2008; Baumeister et al., 2012):[21]

$$(m-1, m-2, \ldots, m-k, m-k-1, \ldots, m-k-1), \text{ for all } 1 \leqslant k \leqslant m$$

- **Averaged Borda** (Dummett, 1997):

$$(m-1, m-2, \ldots, m-k, \frac{m-k-1}{2}, \ldots, \frac{m-k-1}{2}), \text{ for all } 1 \leqslant k \leqslant m$$

Intuitively, the pessimistic Borda rule sees the non-ranked alternatives as placed on the bottom position of a complete ranking, assigning them score 0; the optimistic Borda rule however places those alternatives in the position that immediately follows the last ranked alternative out of the k ones, assigning them score $m-k-1$. The averaged Borda rule assigns to each non-ranked alternative the average of the Borda scores that this alternative could be assigned with if the given top-truncated order were to be extended to a linear order—this is a method that Ackerman et al. (2013) call *bucket averaging*.

Overall, we have reviewed two different—yet equally natural—directions one could follow to generalise the Borda rule on incomplete preferences (and specifically on truncated preferences). We will next see that defining the Borda rule using a domination-based score or a scoring vector can lead to exactly the same outcome. What plays a crucial role here is the particular domain we consider. Specifically, by combining domination-based scores with different domains, we obtain already existing rules (consult Table 3.4).[22] The observations in Table 3.4 are straightforward, except perhaps for the one concerning the averaged Borda rule and the symmetric Borda scores, made explicit in Proposition 3.34.[23]

3.34. Proposition. *The averaged Borda rule for top-truncated preferences (on $\mathcal{D}_1(A)$ or $\mathcal{D}_2(A)$) is the positional scoring rule with a corresponding scoring function based on the symmetric Borda scores.*

Proof. Take the symmetric Borda scores in a top-truncated preference. Divide them by 2 and add the positive constant $\frac{m}{2}$. The new scores are exactly the ones associated

[20]This rule can also be found under the name *truncated Borda* (Ackerman et al., 2013).

[21]This rule has also been referred to as *modified Borda* by Caragiannis et al. (2015).

[22]In Section 3.1, we discussed yet another generalisation of the Borda rule, called *shortsighted Borda*, which is based on the scoring vector $(m-1, m-2, \ldots, m-k, m-k, \ldots, m-k)$. Note that this rule cannot be obtained through domination-based scores, for these particular domains. Indeed, shortsighted Borda is better suited to domains where the unranked alternatives are taken to be incomparable (rather than inferior) to the ranked alternatives.

[23]Proposition 3.34 also holds for larger preference domains, like the one of weak orders, as discussed for instance by Duddy et al. (2016).

Scores \ Domains	$\mathcal{D}_1(A)$	$\mathcal{D}_2(A)$
(a) symmetric	averaged	averaged
(b) domination	pessimistic	pessimistic
(c) weak domination	optimistic	pessimistic
(d) non-domination	optimistic	optimistic

Table 3.4: Borda rules for domains of top-truncated preferences.

with the averaged Borda rule. Since these scores are obtained through an affine transformation of the old ones, the two corresponding rules are equivalent. □

Our three Borda rules often yield different outcomes. For an example, see Figure 3.9.

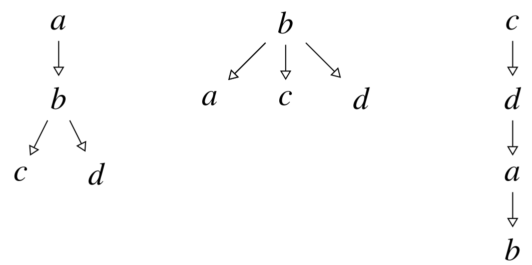


Figure 3.9: Profile where the optimistic, the pessimistic, and the average Borda rule yield different outcomes: $\{b\}$, $\{a, c\}$, and $\{a, b\}$, respectively.

All aforementioned rules belong to a class containing natural generalisations of the Borda rule for top-truncated preferences: the *Borda class* (Figure 3.10).

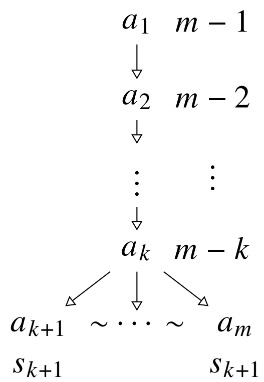


Figure 3.10: A rule in the Borda class on the domain $\mathcal{D}_1(A)$, for $s_{k+1} < m - k$.

A rule for top-truncated preferences is in the Borda class if it is induced by some positional scoring function that gives rise to scoring vectors of the following form, for a number $s_{k+1} < m - k$ and $1 \leqslant k \leqslant m$:

$$(m-1, m-2, \ldots, m-k, s_{k+1}, \ldots, s_{k+1})$$

To sum up this chapter, we have seen various ways in which aggregation rules can be defined when agents hold incomplete opinions, either in the framework of judgment aggregation or in that of preference aggregation and voting. Besides their definitions and the motivation behind those definitions, the discussed rules carry different formal properties, and vary in suitability when we want to reach good, sincere, or correct collective decisions. In the chapters that follow, we will delve deeply in these topics.

Chapter 4

Axiomatic Properties

By now it is apparent that there are plenty of methods one could adopt in order to aggregate (incomplete) individual opinions and produce a collective outcome. But one essential concern regarding all those methods remains unresolved:

> How can we distinguish good from bad procedures of collective decision making, given specific aggregation contexts?

In this chapter, we will follow a simple and effective approach to answer this question. For different aggregation contexts, we will first identify, and then formally define, a number of properties that an appealing aggregation rule should satisfy. Subsequently, we will investigate which ones from the relevant aggregation rules that we have seen in Chapter 3 indeed satisfy the proposed properties. In the most interesting cases, we will be able to show that some of our rules are the unique ones for which a collection of desirable properties simultaneously hold (leading to a characterisation result). In this way, we will provide a strong argument in favour of the selected rules: if we want to ensure that certain normative criteria are going to be fulfilled, then our only option will be to use the particular aggregation rules. On the other hand, we may also find that certain properties can never hold at the same time, no matter what aggregation rule we attempt to use (implying an impossibility result).

What we have so far described as *properties* of aggregation rules are known to the social choice literature as *axioms*.

4.1. Example. Consider a group of three friends, Ann, Bea, and Cal, who need to make a choice about what kind of food to order for dinner, between pizza and burger. Suppose that they are thinking of making a decision by aggregating their preferences as follows: If Ann wants to have pizza, then pizza will be selected; otherwise, they will order what the majority prefers.

This mechanism does not sound very fair in the first place. On the on hand, it assigns some unjustified power to Ann over her friends, and on the other hand, it also distinguishes between the two types of food (since Ann's preference will be followed only in case she is in favour of pizza). In fact, the symmetrical treatment of (i) the *agents* of a group and (*ii*) the *alternatives* about which a collective decision should be reached is ensured by two ubiquitous axioms in social choice theory: *anonymity* and *neutrality*, respectively. Most rules presented in Chapter 3 satisfy these axioms, but the rule of this example does not. △

In this chapter, we will further define and analyse compelling axioms tailored to the aggregation of incomplete opinions. The direction we will take consists of two complementary paths. First, we will discuss the natural generalisation of several axioms that are deemed classical in the field of social choice theory (like *neutrality* or *independence*); second, we will design new axioms that serve their purposes exclusively in domains of incomplete opinions (such an axiom that we will later see is *splitting*). For ease of exposition, we will overload some names of axioms (like the one of *anonymity*), using them across the formal frameworks of judgment aggregation and voting. Of course, whenever applied in a specific section, the name of an axiom will refer to the property defined for the framework within which that section is developed. Axioms with the same name will encode exactly the same idea; on the other hand, general concepts that get further specified within frameworks (like monotonicity in voting vs. monotonicity in judgment aggregation) will be assigned different names.

> Given some possibly infinite sets that encode the superpopulation $N^\star$, the superagenda $\Phi^\star$, and the set of all potential alternatives $A^\star$, our aggregation rules function on all finite subsets of those sets. For the sake of defining axioms in this chapter, we will consider a fixed group of agents $N \subseteq N^\star$ of size n, an agenda $\Phi \subseteq \Phi^\star$, and a set of alternatives $A \subseteq A^\star$. We will say that an aggregation rule satisfies a certain axiom if the relevant definition holds for all such finite N, Φ, and A.

4.1 On Weight Rules

Let us come back to the class of weight rules that we introduced in Section 3.1. Supposing that we are going to use some rule from this class, we wonder whether there are natural axioms that can discriminate between all available options. We conduct our analysis for the framework of judgment aggregation, as well as for preference aggregation and voting. Although the core ideas behind our results remain the same across all frameworks, the details in several proofs need to be accounted for separately. We discuss the technical correspondence between our various findings as we proceed.

We present three kinds of axioms that encode attractive normative requirements for aggregation rules that operate on possibly incomplete opinions:

- *Majoritarianism.* When deciding between accepting a proposition φ and its negation $\neg\varphi$, or when contemplating how to rank two alternatives a and b, it is appealing to follow the opinion of the majority of agents. But, as is well known, doing so can lead to logically inconsistent outcomes and cyclic preferences. We therefore propose three weakened forms of majoritarianism that avoid this conflict between logical consistency or acyclicity on the one hand, and responsiveness to the will of the majority on the other hand.

- *Splitting.* Suppose that a subgroup of all agents observe that their (disjoint) individual opinions are compatible with each other (meaning that the union of these opinions constitutes a new admissible opinion). Then, they might consider submitting that union rather than their own individual opinions. It seems desirable to use an aggregation rule for which the outcome never changes when a subgroup chooses to make such a move—we consider axioms of varying strength based on this fundamental idea.

- *Quality over quantity.* Consider some agents that hold opinions on a small number of topics. We may think that these agents have reflected on those topics more carefully than other agents who have opinions on a large number of topics; we may also regard it a matter of fairness to either give someone much influence over a few topics or little influence over many, but not much influence over many topics. We propose an axiom that takes an extreme position on this matter and requires that a single "quality" agent with an opinion on few topics is always more powerful than any number of agents with the same opinions on a strictly larger set of topics.

Our main results are characterisations of weight rules that satisfy the axioms sketched above. Under appropriate conditions, imposing majoritarianism demands constant weights. A (weak form of the) splitting axiom causes weights to be inversely proportional to the number of topics on which an agent reports an opinion (which is the case for the equal-and-even weight rule). Finally, the quality-over-quantity axiom forces the rule to be the upward-lexicographic one, making decisions by ordering the agents lexicographicaly, according to the number of topics upon which they express an opinion.

4.1.1 Judgment Aggregation

In this section, we discuss a number of reasonable features of weight rules for incomplete individual judgments, and we formalise them in terms of axioms. Then, we demonstrate how specific weight rules are charaterised via these axioms.

Axioms

To begin with, making a collective decision based on the opinion of the majority on each issue of the agenda separately is commonly considered as a desirable attribute of an aggregation rule. However, we know that majorities easily lead to inconsistencies when logical interconnections between the propositions are involved (we refer, for instance, to the *doctrinal paradox* by Kornhauser and Sager, 1993). On the other hand, there clearly are situations where the role of logic is not critical for the decision-making process, while the judgment of the majority is still considered a reliable indication concerning the judgment of the whole group.

4.2. Example. Suppose that the jury of a court has to deliver a verdict on two distinct cases, deciding whether the two defendants are guilty (propositions φ_1 and φ_2). In this scenario, no matter what the verdict on φ_1 is, the jury can independently decide in favour of φ_2 or of $\neg\varphi_2$. Formally, for the agenda $\Phi = \{\varphi_1, \varphi_2, \neg\varphi_1, \neg\varphi_2\}$ and for a jury of size n, it holds that φ_2 is logically independent of $\boldsymbol{J}$ for every profile $\boldsymbol{J} \in \mathcal{J}(\Phi)^n$. Now, imagine that the jury consists of five members, three of which find that the defendant of the second case is guilty. This means that a strict majority of the jury members accept proposition φ_2. The axiom of *forward majoritarianism* says exactly that this majority ought to be respected by an aggregation rule; regardless of what the jury decides about the first case, any collective judgment should contain proposition φ_2. Then, suppose that the judge of the court does not have access to the individual judgments of the jury members, but she can observe the outcome of an aggregation rule, where every collective judgment (in case there are more than one) contains φ_2. The judge cannot avoid but announce the defendant of the second case guilty, without knowing how many out of the five jury members agree with this verdict. The axiom of *backward majoritarianism* provides the judge with the guarantee that a strict majority of the jury members have deemed the defendant guilty. △

For a profile $\boldsymbol{J} \in \mathcal{J}(\Phi)^n$, let us define the *simple-majority set*, that is, the set of all propositions that have more advocates than adversaries in $\boldsymbol{J}$:

$$m(\boldsymbol{J}) = \{\varphi \in \Phi \mid n^{\boldsymbol{J}}_{\varphi} > n^{\boldsymbol{J}}_{\neg\varphi}\}$$

> One could also work with the *absolute-majority set* $\{\varphi \in \Phi \mid n^{\boldsymbol{J}}_{\varphi} > \frac{n}{2}\}$ and obtain the same technical results. By adopting this definition, our axioms of forward majoritarianism and of general majoritarianism would become weaker, while backward majoritarianism would be logically stronger.

Then, we say that a proposition φ is *logically independent* of a profile $\boldsymbol{J}$ whenever φ is logically independent of each consistent subset of $\{\psi, \neg\psi \mid \psi \in \bigcup_{i \in N} J_i\} \setminus \{\varphi, \neg\varphi\}$.[1] Intuitively, if a proposition is independent of a profile, then we can flip it in every agent's individual judgment (i.e., replace φ with $\neg\varphi$ and vice versa) without damaging logical consistency. See Figure 4.1 for an example.

Let us now formulate three concrete axioms that encode majoritarian ideals.

Axiom: ***Forward Majoritarianism.*** *For every profile $\boldsymbol{J} \in \mathcal{J}(\Phi)^n$ and proposition $\varphi \in \Phi$ that is logically independent of $\boldsymbol{J}$, it holds that*

$$\textit{if } \varphi \in m(\boldsymbol{J}), \textit{ then } \varphi \in J \textit{ for all } J \in F(\boldsymbol{J})$$

[1]Recall that a proposition φ is logically independent of a set of propositions Y if both $Y \cup \{\varphi\}$ and $Y \cup \{\neg\varphi\}$ are logically consistent.

$J_1 = \{p, q\}$ $\quad\quad\quad\quad$ $J_1 = \{p, q, p \wedge q\}$

$J_2 = \{p, \neg q, \neg(p \wedge q)\}$ $\quad\quad\quad\quad$ $J_2 = \{p, \neg q, \neg(p \wedge q)\}$

Figure 4.1: On the left, p is independent of the profile; on the right, it is not.

Axiom: ***Backward Majoritarianism.*** *For every profile $\boldsymbol{J} \in \mathcal{J}(\Phi)^n$ and proposition $\varphi \in \Phi$ that is logically independent of $\boldsymbol{J}$, the following holds:*

$$\textit{if } \varphi \in J \textit{ for all } J \in F(\boldsymbol{J}), \textit{ then } \varphi \in m(\boldsymbol{J})$$

Regarding the general matter of majority judgments conflicting with logical constraints, a notion of minimal divergence from the majority outcome based on set-inclusion can be defined: For a set of propositions $Y \subseteq \Phi$, $Y' \subseteq Y$ is a maximal consistent subset of Y if and only if Y' is consistent and there exists no other consistent set Y'' such that $Y' \subset Y'' \subseteq Y$. The set of maximal consistent subsets of Y is denoted by $\max(Y, \subseteq)$. Following Nehring et al. (2014), we call $Con(\boldsymbol{J}) = \{Y' \mid Y'$ complete and consistent, and $Y' \supseteq Y$, for some $Y \in \max(m(\boldsymbol{J}), \subseteq)\}$ the *Condorcet set*.

General majoritarianism ensures that the aggregation rule induces only judgment sets that extend (possibly in an inconsistent way) judgment sets in the Condorcet set.[2]

Axiom: ***General Majoritarianism.*** *For every profile $\boldsymbol{J} \in \mathcal{J}(\Phi)^n$ and every judgment set $J \in F(\boldsymbol{J})$, there exists a judgment set $J' \in Con(\boldsymbol{J})$ such that $J' \subseteq J$.*

By inspecting our definitions, we can see that, as far as all aggregation rules are concerned, general majoritarianism logically implies forward majoritarianism and both these axioms are independent of backward majoritarianism. However, we will later show that the relations between these axioms differ when we focus on weight rules.

We illustrate the intuition bared by our next axioms with an example.

4.3. Example. A group of doctors has to make a decision concerning the treatment of a patient. They must decide whether treatment 1 is suitable (proposition φ_1), similarly for treatment 2 (proposition φ_2), but also whether treatment 1 and 2 can be combined (proposition $\varphi_1 \wedge \varphi_2$), etc. Suppose that the doctors have different specialisations (or that they choose to spend their limited time investigating different issues regarding the decision they have to make). This would result in a meeting where these doctors submit judgments on different propositions. Suppose now that a subset of the doctors have the chance to eat lunch together before the meeting, during which they discuss their patient. They are then able to share the information they collected individually and the judgments at which they arrived. This subset of the doctors realise that all their judgments together form a consistent set of propositions $\{\varphi_1, \ldots, \varphi_\lambda\}$. Trusting

[2]If the aggregation rule induces only complete and consistent collective judgments, general majoritarianism means that $F(\boldsymbol{J}) \subseteq Con(\boldsymbol{J})$ for every profile $\boldsymbol{J} \in \mathcal{J}(\Phi)^n$.

their colleagues, the doctors in this subset may report the whole set of propositions in the meeting, instead of the more restricted judgments they held before the discussion. We would like the aggregation rule employed by the group to reach the same outcome independently of whether some of the doctors choose to combine their individual judgments prior to aggregation. △

We present three axioms that enforce exactly the propoerty we hinted at in Exampe 4.3 in various degrees, ordered from stronger to weaker:

Axiom: ***Arbitrary Splitting.*** *For every profile $\boldsymbol{J} \in \mathcal{J}(\Phi)^n$ and subgroup $\emptyset \neq N' \subseteq N$ of agents with pairwise disjoint and mutually consistent judgment sets, we have $F(\boldsymbol{J}) = F(\boldsymbol{J}')$, where $\boldsymbol{J}'$ arises from $\boldsymbol{J}$ by replacing the judgment set of each member of N' by the union $\bigcup_{i \in N'} J_i$.*

Axiom: ***Equal Splitting.*** *For every profile $\boldsymbol{J} \in \mathcal{J}(\Phi)^n$ and subgroup $\emptyset \neq N' \subseteq N$ of agents whose judgment sets are pairwise disjoint, mutually consistent, and of equal size, we have $F(\boldsymbol{J}) = F(\boldsymbol{J}')$, where $\boldsymbol{J}'$ arises from $\boldsymbol{J}$ by replacing the judgment set of each member of N' by the union $\bigcup_{i \in N'} J_i$.*

Axiom: ***Single Splitting.*** *For every profile $\boldsymbol{J} \in \mathcal{J}(\Phi)^n$ and subgroup $\emptyset \neq N' \subseteq N$ of agents whose judgment sets are pairwise disjoint, mutually consistent, and singletons, we have $F(\boldsymbol{J}) = F(\boldsymbol{J}')$, where $\boldsymbol{J}'$ arises from $\boldsymbol{J}$ by replacing the judgment set of each member of N' by the union $\bigcup_{i \in N'} J_i$.*

The next property is particularly desirable when smaller judgment sets are significantly more important to the collective decision than larger ones. For instance, in case the agents have a fixed and limited amount of energy/time/cognitive effort at their disposal, then one could expect that small judgment sets are more well-thought-out than large ones, and hence of greater value to the group. We restrict attention to profiles where only two judgments—which differ in size—are reported. The axiom of *quality-over-quantity* states that the collective judgment ought to always agree with the smaller individual judgment set, independently of how many agents adopt each opinion:[3]

Axiom: ***Quality-over-Quantity.*** *Consider two arbitrary judgment sets $\emptyset \neq J', J'' \in \mathcal{J}(\Phi)$ with $|J'| < |J''|$. For every profile $\boldsymbol{J} \in \mathcal{J}(\Phi)^n$ where a subgroup $\emptyset \neq N' \subset N$ of agents declare judgment J', another subgroup $\emptyset \neq N'' \subset N$ of agents declare judgment J'', and the rest of the agents submit an empty set, it holds that $J' \subseteq J$ for all $J \in F(\boldsymbol{J})$.*

[3]The property of quality-over-quantity may be judged as weak or strong, depending on different considerations. On the one hand, the axiom's scope (i.e., the profiles that its antecedent refers to) may categorise it as weak, since it only mentions cases where just two distinct judgment sets are reported. On the other hand, the axiom may also be considered strong, because it clearly prioritises small judgment sets over large ones, not allowing for any compromise.

Characterisation Results

Using the axioms that we have presented so far, we will provide axiomatic characterisations for aggregation rules within the class of weight rules.

It will be important that the agendas containing the issues at stake have a minimum level of variety. To that end, we define the notion of *unconstrainedness*.

For $\lambda \in \mathbb{N}$, we say that a superagenda $\Phi^\star$ is *λ-constrained* if for every logically consistent subset $J \subseteq \Phi^\star$ with $|J| = \lambda$, the following holds for all $\varphi \in J$:

$$J \setminus \{\varphi\} \vDash \varphi$$

In words, λ-constrainedness means that every proposition in a consistent set of size λ is implied by all the other propositions in that set together. Suppose that $\Phi^\star$ is not λ-constrained for some λ. Then, there exist a consistent judgment set J of size λ and a proposition $\varphi \in J$ such that $(J \setminus \varphi) \cup \{\neg\varphi\}$ is logically consistent too, meaning that we can replace φ with its negation without damaging J's consistency. We will often exploit this attribute in our proofs.

A superagenda $\Phi^\star$ is *unconstrained* if there is no $\lambda \leqslant \frac{|\Phi^\star|}{2}$ for which $\Phi^\star$ is λ-constrained. Unconstrainedness is a very weak property. In particular, every superagenda becomes unconstrained if we add to it a new propositional variable and its negation. Moreover, most agendas used in common judgment aggregation applications (like those containing a conjunctive or disjunctive agenda with at least two premises (Dietrich and List, 2007c), a preference agenda (Dietrich and List, 2007a) or a budget agenda (Dietrich and List, 2010b)) are unconstrained.

We start by investigating the three majoritarian properties. For a finite superpopulation, we show that each one of the axioms of general majoritarianism and of forward majoritarianism characterises a family of rules with weights that are “close enough" (Proposition 4.4), while if the superpopulation is infinite we obtain an axiomatisation of the constant weight rule (Theorem 4.5).

4.4. Proposition. *For any unconstrained superagenda $\Phi^\star$, finite superpopulation $N^\star$, and weight function w, the weight rule F_w satisfies general majoritarianism if and only if it satisfies forward majoritarianism if and only if $\frac{w_\lambda}{w_{\lambda'}} < \frac{k}{k-1}$ for all $\lambda, \lambda' \in \mathbb{N}$ with $\lambda, \lambda' \leqslant \frac{|\Phi^\star|}{2}$, where $k = \lceil\frac{|N^\star|}{2}\rceil$.*

Proof. We will explain the proof only for the property of general majoritarianism, since the proof for forward majoritarianism is similar. Given an unconstrained superagenda $\Phi^\star$ and a finite superpopulation $N^\star$, we address the case for an even number $|N^\star| = 2k$ (when $|N^\star|$ is odd the proof is totally analogous). Consider an arbitrary weight function w and its induced weight rule F_w.

Suppose that F_w satisfies general majoritarianism and consider two arbitrary $\lambda, \lambda' \in \mathbb{N}$. Then, if $w_\lambda = w_{\lambda'}$, we have that $\frac{w_\lambda}{w_{\lambda'}} < \frac{k}{k-1}$. So, assume that $w_\lambda > w_{\lambda'}$ and $\lambda > \lambda'$ (if $\lambda < \lambda'$ the proof is analogous). Since $\Phi^\star$ is unconstrained, there exists a consistent set

$J = \{\varphi_1, \varphi_2, \ldots, \varphi_\lambda\} \subseteq \Phi^\star$ such that $J' = \{\neg\varphi_1, \varphi_2, \ldots, \varphi_\lambda\}$ is also consistent. Take J'' to be a consistent subset of J' with $\neg\varphi_1 \in J''$ and $|J''| = \lambda'$. Consider the following profile, corresponding to a group of $|N^\star|$ agents:

$$\boldsymbol{J} = (\emptyset, \underbrace{J, \ldots, J}_{k-1}, \underbrace{J'', \ldots, J''}_{k})$$

Since $n^{\boldsymbol{J}}_{\neg\varphi_1} > n^{\boldsymbol{J}}_{\varphi_1}$ and $J'' \subset J' = (J \setminus \{\varphi_1\}) \cup \{\neg\varphi_1\}$, it is true that $Con(\boldsymbol{J}) = \{J'\}$. Then, F_w satisfying general majoritarianism implies that $F_w(\boldsymbol{J}) = \{J'\}$, so φ_1 must weigh strictly less than $\neg\varphi_1$ in $\boldsymbol{J}$. That is, $(k-1) \cdot w_\lambda < k \cdot w_{\lambda'}$ or $\frac{w_\lambda}{w_{\lambda'}} < \frac{k}{k-1}$.

For the other direction we will show that if F_w does not satisfy general majoritarianism, then $\frac{w_\lambda}{w_{\lambda'}} \geqslant \frac{k}{k-1}$ for some λ, λ'.

We start with assuming that F_w does not satisfy general majoritarianism. Since F_w—being a weight rule—produces only complete and consistent collective judgments, failing general majoritarianism means that $F_w(\boldsymbol{J}) \nsubseteq Con(\boldsymbol{J})$ for some agenda $\Phi \subseteq \Phi^\star$ and some profile $\boldsymbol{J} \in \mathcal{J}(\Phi)^n$ with $n \leqslant |N^\star|$. This means that there exists a judgment set $J \in F_w(\boldsymbol{J})$ such that $J \notin Con(\boldsymbol{J})$. The fact that $J \notin Con(\boldsymbol{J})$ implies that there are some propositions $\varphi_1, \ldots, \varphi_\ell \in J$ such that $n^{\boldsymbol{J}}_{\neg\varphi_j} > n^{\boldsymbol{J}}_{\varphi_j}$ for all $j \in \{1, \ldots, \ell\}$, and $J' = (J \setminus \{\varphi_1, \ldots, \varphi_\ell\}) \cup \{\neg\varphi_1, \ldots, \neg\varphi_\ell\}$ is consistent. Because of $J \in F_w(\boldsymbol{J})$, we have the following:

$$\begin{aligned}
& \sum_{i \in N} w_{|J_i|} \cdot |J \cap J_i| \geqslant \sum_{i \in N} w_{|J_i|} \cdot |J' \cap J_i| \\
\Rightarrow \quad & \sum_{j \in \{1,\ldots,\ell\}} \sum_{i \in N^{\boldsymbol{J}}_{\varphi_j}} w_{|J_i|} \geqslant \sum_{j \in \{1,\ldots,\ell\}} \sum_{i \in N^{\boldsymbol{J}}_{\neg\varphi_j}} w_{|J_i|} \\
\Rightarrow \quad & (n^{\boldsymbol{J}}_{\varphi_1} + \ldots + n^{\boldsymbol{J}}_{\varphi_\ell}) \cdot \max_{|J_i|} w_{|J_i|} \geqslant (n^{\boldsymbol{J}}_{\neg\varphi} + \ldots + n^{\boldsymbol{J}}_{\neg\varphi_\ell}) \cdot \min_{|J_i|} w_{|J_i|}
\end{aligned}$$

Since $n^{\boldsymbol{J}}_{\varphi_1} + \ldots + n^{\boldsymbol{J}}_{\varphi_\ell} > 0$, the following holds:

$$\frac{n^{\boldsymbol{J}}_{\neg\varphi_1} + \ldots + n^{\boldsymbol{J}}_{\neg\varphi_\ell}}{n^{\boldsymbol{J}}_{\varphi_1} + \ldots + n^{\boldsymbol{J}}_{\varphi_\ell}} \geqslant \frac{n^{\boldsymbol{J}}_{\varphi_1} + \ldots + n^{\boldsymbol{J}}_{\varphi_\ell} + \ell}{n^{\boldsymbol{J}}_{\varphi_1} + \ldots + n^{\boldsymbol{J}}_{\varphi_\ell}} = 1 + \frac{\ell}{n^{\boldsymbol{J}}_{\varphi_1} + \ldots + n^{\boldsymbol{J}}_{\varphi_\ell}} \geqslant 1 + \frac{\ell}{\ell(k-1)} = \frac{k}{k-1}.$$

We conclude that $(k-1) \cdot \max_{|J_i|} w_{|J_i|} \geqslant k \cdot \min_\lambda w_\lambda$; this means that for some λ, λ', it is the case that $\frac{w_\lambda}{w_{\lambda'}} \geqslant \frac{k}{k-1}$. □

4.5. THEOREM. *For any unconstrained superagenda $\Phi^\star$ and infinite superpopulation $N^\star$, a weight rule satisfies general majoritarianism if and only if it satisfies forward majoritarianism if and only if it is the constant weight rule F_c.*

PROOF. If general (forward) majoritarianism holds for an infinite superpopulation $N^\star$, then it has to hold for every finite group of agents. As in the proof of Proposition 4.4, we must have that $\frac{w_\lambda}{w_{\lambda'}} < \frac{k}{k-1}$ for all $k \in \mathbb{N}$. This only holds if $w_\lambda = w_{\lambda'}$ for all λ, λ'. □

In case, however, we wish to make no assumptions regarding the superpopulation, the constant weight rule F_c is directly characterised by backward majoritarianism.

4.6. Theorem. *For any superpopulation and unconstrained superagenda, a weight rule satisfies backward majoritarianism if and only if it is the constant weight rule F_c.*

Proof. Consider an unconstrained superagenda $\Phi^\star$ and a superpopulation $N^\star$. For the first direction we take an arbitrary profile $\boldsymbol{J} = (J_1, \dots, J_n) \in \mathcal{J}(\Phi)^n$ for some $\Phi \subseteq \Phi^\star$ and some $n \leqslant |N^\star|$ and a proposition $\varphi \in \Phi$ logically independent of $\boldsymbol{J}$. The independence assumption for φ implies that the constant weight rule F^c with $c > 0$ will simply count how many agents agree with the propositions φ and $\neg\varphi$ in the profile $\boldsymbol{J}$, and proceed as follows: if $n^{\boldsymbol{J}}_\varphi = n^{\boldsymbol{J}}_{\neg\varphi}$, then both φ and $\neg\varphi$ will belong to some collective judgment set in $F^c(\boldsymbol{J})$; if $n^{\boldsymbol{J}}_\varphi > n^{\boldsymbol{J}}_{\sim\varphi}$, then only φ will belong to some collective judgment set in $F^c(\boldsymbol{J})$; and if $n^{\boldsymbol{J}}_\varphi < n^{\boldsymbol{J}}_{\neg\varphi}$, then only $\neg\varphi$ will belong to some collective judgment set in $F^c(\boldsymbol{J})$. This means that if $\varphi \in J$ for all $J \in F^c(\boldsymbol{J})$, then $n^{\boldsymbol{J}}_\varphi > n^{\boldsymbol{J}}_{\sim\varphi}$, and $\varphi \in m(\boldsymbol{J})$. So backward majoritarianism holds.

For the other direction, we need to show that if backward majoritarianism is satisfied by some weight rule F_w induced by a neutral weight function w, then it must hold that $w_\lambda = w_{\lambda'}$ for all λ, λ'. We proceed with proving the contrapositive. Take a weight function w such that $w_\lambda \neq w_{\lambda'}$ for some $\lambda \neq \lambda'$. Suppose that $w_\lambda > w_{\lambda'}$ and $\lambda > \lambda'$ (if $\lambda < \lambda'$ the proof is analogous). Then, since $\Phi^\star$ is unconstrained, there is some consistent judgment set $J \subseteq \Phi^\star$ of size λ such that $\varphi \in J$ and $J' = (J \setminus \{\varphi\}) \cup \{\neg\varphi\}$ is also consistent. Consider the agenda $\Phi \subseteq \Phi^\star$ that contains all propositions in J and their negations, and take $J'' \subset J'$ such that $\neg\varphi \in J''$ and $|J''| = \lambda'$. Then, we construct the profile $\boldsymbol{J} = (\emptyset, \dots, \emptyset, J, J'')$, where by definition, φ is logically independent of $\boldsymbol{J}$. Since $w_{|J''|} = w_{\lambda'} < w_\lambda = w_{|J|}$, we will have that $\varphi \in J$ for all $J \in F_w(\boldsymbol{J})$. However, $n^{\boldsymbol{J}}_\varphi = n^{\boldsymbol{J}}_{\neg\varphi} = 1$, thus $\varphi \notin m(\boldsymbol{J})$. We conclude that backward majoritarianism fails. □

Our results highlight an interesting fact about the logical relations between the three majoritarian axioms, when restricting attention to the class of weight rules. Forward and of general majoritarianism are now logically equivalent (even though in general the latter is stronger than the former), and logically weaker than backward majoritarianism (which in general is logically independent of both).

We proceed with exploring the splitting axioms with respsect to weight rules. For a sufficiently large superpopulation, we show that the weakest axiom of the three, namely single splitting, characterises the equal-and-even weight rule (Theorem 4.7), and by Theorem 4.8 the same holds for equal splitting. Moreover, arbitrary splitting is proven to be too strong: it is not satisfied by any weight rule (Proposition 4.9).

The characterisation results that rely on splitting would fail without a sufficiently large superpopulation; dropping this assumption would lead to a characterisation of some uninspiring class of rules instead of a single rule.

4.7. Theorem. *For any unconstrained superagenda $\Phi^\star$ and superpopulation $N^\star$ with $|N^\star| \geqslant \frac{|\Phi^\star|}{2} + 1$, a weight rule satisfies single splitting if and only if it is the equal-and-even weight rule F_{ee}.*

Proof. Consider given an unconstrained superagenda $\Phi^\star$ and a superpopulation $N^\star$ with $|N^\star| \geqslant \frac{|\Phi^\star|}{2} + 1$. For the first direction, we consider an arbitrary instance of the single splitting axiom: an agenda $\Phi \subseteq \Phi^\star$, a group N of $n \leqslant |N^\star|$ agents, a non-empty subgroup $N' \subseteq N$ of agents whose judgment sets are pairwise disjoint, mutually consistent, and singleton, and two profiles $\boldsymbol{J} = (J_1, \ldots, J_n)$ and $\boldsymbol{J}'$, where $\boldsymbol{J}'$ arises from $\boldsymbol{J}$ by replacing the judgment set of each member of N' by the union $\bigcup_{i \in N'} J_i$. By the definition of the weights for the rule F_{ee} we know that every proposition in $\bigcup_{i \in N'} J_i$ weighs exactly the same in the profiles $\boldsymbol{J}$ and $\boldsymbol{J}'$, because $|J_i| = 1$ for every $i \in N'$, and $w_\lambda = \frac{w_1}{\lambda}$ for every λ. The same holds for all propositions that are not in $\bigcup_{i \in N'} J_i$ too, since they trivially appear in exactly the same judgment sets in both profiles $\boldsymbol{J}$ and $\boldsymbol{J}'$. This means that $F_{ee}(\boldsymbol{J}) = F_{ee}(\boldsymbol{J}')$.

For the other direction, consider an arbitrary $\lambda \leqslant |N^\star| - 1$ (note that we do not need to consider larger values for λ because $|N^\star| \geqslant \frac{|\Phi^\star|}{2} + 1$). Since $\Phi^\star$ is unconstrained, we take a consistent subset $J = \{\varphi_1, \varphi_2, \ldots, \varphi_\lambda\}$ such that $\{\neg\varphi_1, \varphi_2, \ldots, \varphi_\lambda\}$ is also consistent. Consider the agenda Φ that contains all propositions in J and their negations and take two profiles as follows:

$$\boldsymbol{J} = (\{\neg\varphi_1\}, \{\varphi_1\}, \{\varphi_2\}, \ldots, \{\varphi_\lambda\})$$
$$\boldsymbol{J}' = (\{\neg\varphi_1\}, \{\varphi_1, \varphi_2, \ldots, \varphi_\lambda\}, \{\varphi_1, \varphi_2, \ldots, \varphi_\lambda\}, \ldots, \{\varphi_1, \varphi_2, \ldots, \varphi_\lambda\})$$

Note that this can be done because $\lambda \leqslant |N^\star| - 1$. Then, if a weight rule F_w induced by a weight function w satisfies the single splitting axiom, it must be the case that $F_w(\boldsymbol{J}) = F_w(\boldsymbol{J}')$. So the weights of φ_1 and $\neg\varphi_1$ in profile $\boldsymbol{J}'$ should coincide, that is, $w_1 = \lambda \cdot w_\lambda$. Hence, by Definition 3.11, F_w is the equal-and-even weight rule. □

4.8. Theorem. *For any unconstrained superagenda $\Phi^\star$ and any superpopulation $N^\star$ with $|N^\star| \geqslant \frac{|\Phi^\star|}{2} + 1$, a weight rule satisfies equal splitting if and only if it is the equal-and-even weight rule F_{ee}.*

Proof. Trivially, if a weight rule F_w satisfies the equal splitting axiom, then it satisfies the single splitting axiom too, so Theorem 4.7 implies that F_w must be the equal-and-even weight rule. For the other direction it is easy to see that for all weight of the form $w_\lambda = \frac{w_1}{\lambda}$ (which induce the equal-and-even weight rule by Definition 3.11), the equal splitting axiom is satisfied (note, however, that the arbitrary splitting axiom is not). □

> Recalling the formulations of the splitting axioms, the reader may be curious to know how essential for our characterisation results the assumption of disjointness of the individual judgments is. The answer is “very"—without it, the splitting properties would fail for all weight rules.

Finally, we prove an impossibility result concerning the arbitrary splitting axiom and the class of weight rules.

4.9. Proposition. *For any unconstrained superagenda* $\Phi^\star$ *and superpopulation* $N^\star$ *with* $|N^\star| > 2$ *and* $|\Phi^\star| > 2$*, there exists no weight rule that satisfies arbitrary splitting.*

Proof. For the sake of contradiction, assume that there exists some weight function w for which the induced weight rule F_w satisfies the arbitrary splitting axiom. Then, F_w should clearly also satisfy the single splitting axiom. Hence, by Theorem 4.7, we should have that $w_\lambda = \frac{w_1}{\lambda}$ for every λ. However, we will now show that if this is the case, then F_w can in fact not satisfy the arbitrary splitting axiom, which stands in direct contradiction to our assumption.

Consider three propositions $\varphi_1, \varphi_2, \varphi_3 \in \Phi^\star$ such that both sets $\{\varphi_1, \varphi_2, \varphi_3\}$ and $\{\neg\varphi_1, \varphi_2, \varphi_3\}$ are consistent (which is possible due to unconstrainedness), and take the agenda $\Phi = \{\varphi_1, \varphi_2, \varphi_3, \neg\varphi_1, \neg\varphi_2, \neg\varphi_3\}$. Then, consider two profiles

$$\begin{aligned} \boldsymbol{J} &= (\{\neg\varphi_1\}, \{\varphi_1\}, \{\varphi_2, \varphi_3\}) \\ \boldsymbol{J}' &= (\{\neg\varphi_1\}, \{\varphi_1, \varphi_2, \varphi_3\}, \{\varphi_1, \varphi_2, \varphi_3\}) \end{aligned}$$

Since both φ_1 and $\neg\varphi_1$ have weight w_1 in profile $\boldsymbol{J}$, it holds that both φ_1 and $\neg\varphi_1$ belong to some judgment set in $F_w(\boldsymbol{J})$. Now, note that $w_{|\{\varphi_1,\varphi_2,\varphi_3\}|} = \frac{w_1}{3}$ (this is true given that $|N^\star| \geqslant 4$ and thus Theorem 4.7 holds). So, φ_1 scores $\frac{2w_1}{3} < w_1$ in $\boldsymbol{J}'$. This means that for every judgment set in $F_w(\boldsymbol{J}')$, only $\neg\varphi_1$ but not φ_1 will belong to that judgment set. To conclude, $F_w(\boldsymbol{J}) \neq F_w(\boldsymbol{J}')$, which violates the arbitrary splitting axiom. □

> An alternative idea in the spirit of splitting would be to require that the outcome of the aggregation process does not change when a subgroup of the agents combine their individual judgments and submit the consistent union a single time (instead of all members of the subgroup submitting that union). This axiom would characterise the constant weight rule: The weight of a proposition will not depend on the size of the judgment set to which it belongs.

Our results illuminate a straightforward separation between weight rules that are majoritarian on the one hand, and those that satisfy the splitting axioms on the other, giving rise to an impossibility result (stated for an infinite superpopulation for simplicity).

4.10. Corollary. *For any unconstrained superagenda* $\Phi^\star$ *and any infinite superpopultation* $N^\star$*, there is no weight rule that satisfies some form of majoritarianism together with some form of splitting.*

As the reader may have already predicted, we will finally prove that the axiom of quality-over-quantity characterises the upward-lexicographic rule.

> The axiomatisation of Theorem 4.11 holds for a finite superagenda $\Phi^\star$ too, under the condition that for every $\lambda \leqslant \frac{|\Phi^\star|}{2}$ there exists a set of propositional variables $\{p_1, \ldots, p_\lambda\} \subseteq \Phi^\star$ such that $\neg p_1 \wedge \cdots \wedge \neg p_\lambda \in \Phi^\star$. However, if we take an infinite superpopulation $N^\star$ instead of a finite one, then the characterisation result does not hold. Specifically, the upward-lexicographic rule cannot be defined as a weight rule anymore, and the quality-over-quantity axiom is not satisfied by any weight rule.

4.11. Theorem. *For any superagenda that is closed under conjunction of literals and contains all propositional variables ,and for any finite superpopulation, a weight rule satisfies quality-over-quantity if and only if it is the upward-lexicographic rule F_{ulex}.*

Proof. It is easy to verify that the upward-lexicographic rule satisfies the quality-over-quantity axiom. For the other direction, consider any finite superpopulation $N^\star$ and any superagenda $\Phi^\star$ of the required kind. Let F_w be a weight rule induced by w.

Consider an arbitrary $\lambda > 1$. We are going to establish a lowe r bound on the ratio $\frac{w_{\lambda-1}}{w_\lambda}$ based on quality-over-quantity. As $\Phi^\star$ is closed under conjunction of literals and contains all propositional variables, we can construct an agenda Φ with $p_1, \ldots, p_{2\lambda-2} \in \Phi$ and $\neg p_1 \wedge \cdots \wedge \neg p_\lambda \in \Phi$. We then use the following two judgment sets J and J', in order to construct the profile $\boldsymbol{J} = (\underbrace{J, \ldots, J}_{|N^\star|-1}, J')$:

$$J = \{p_1, \ldots, p_\lambda\}$$
$$J' = \{\neg p_1 \wedge \cdots \wedge \neg p_\lambda, p_{\lambda+1}, \ldots, p_{2\lambda-2}\}$$

As $|J| = \lambda$ and $|J'| = \lambda - 1$, the quality-over-quantity axiom requires that $J' \subseteq J''$ for all $J'' \in F_w(\boldsymbol{J})$. Given that the propositions $p_{\lambda+1}, \ldots, p_{2\lambda-2}$ are logically independent of the profile $\boldsymbol{J}$, having that $J' \subseteq J''$ for all $J'' \in F_w(\boldsymbol{J})$ means that accepting $\neg p_1 \wedge \cdots \wedge \neg p_\lambda$ must have yielded a higher total weight than accepting all of $p_1, \ldots, p_\lambda$, i.e., we must have that $w_{\boldsymbol{J}}(\{\neg p_1 \wedge \cdots \wedge \neg p_\lambda\}) > w_{\boldsymbol{J}}(J)$. But this is equivalent to $1 \cdot 1 \cdot w_{\lambda-1} > (|N^\star| - 1) \cdot \lambda \cdot w_\lambda$, which in turn is equivalent to $\frac{w_{\lambda-1}}{w_\lambda} > \lambda \cdot (|N^\star| - 1)$. The claim follows from the property of F_{ulex} given by Lemma 3.5 in Chapter 3. □

4.1.2 Voting and Preference Aggregation

In this section, we investigate axioms that are counterparts of those that of Section 4.1.1, but that this time apply to weight rules in the context of voting and preference aggregation. Once more, it is important to remember that we can interpret a relative ranking between two alternatives a and b as accepting a proposition of the form "a is above b". Hence, several of our formalisations and results will follow from the analogous ones in judgment aggregation. As in Section 3.1.2 where weight rules were defined, here we also work with strict acyclic preferences P, which are not necessarily transitive.

> For all results of this section, we have to assume an infinite set of potential alternatives. There is a main difference with the framework of judgment aggregation, where we did not always need to assume an infinite superagenda. In judgment aggregation, the largest λ for which we need to define a weight w_λ is the number of all positive propositions in the superagenda, i.e., $\frac{|\Phi^\star|}{2}$. In order to prove statements for this number, it suffices to consider an agenda $\Phi \subseteq \Phi^\star$ such that $|\Phi| = |\Phi^\star|$. But in preference aggregation, the largest λ for which we need to define a weight w_λ is the number of all pairs of alternatives from the set of all potential alternatives, i.e., $\frac{|A^\star|(|A^\star|-1)}{2}$. In order to prove statements for this number (specifically for rules that output preferences of winner type), we may need to have a set of alternatives $A \subseteq A^\star$ such that $|A| = \frac{|A^\star|(|A^\star|-1)}{2}$, which is not possible if $|A^\star|$ is finite. Although we do not construct such a large set of alternatives in our proofs, employing a set A with $|A| > \lambda$ to prove a statement about λ implies that we need an infinite $A^\star$.

Majoritarianism

We start with the elementary majority concept. In voting scenarios, respecting the preferences of the majority is frequently taken on board as the ideal goal of an aggregation process, although it easily leads to preference cycles (de Condorcet, 1785). For a profile $\boldsymbol{P} \in \mathcal{P}(A)^n$, we define the *simple-majority set* as follows:

$$m(\boldsymbol{P}) = \{(a, b) \in A{\times}A \mid n^{\boldsymbol{P}}_{ab} > n^{\boldsymbol{P}}_{ba}\}$$

There are clear situations where pairs of alternatives cannot cause a Condorcet-type paradox on a profile. Such is the case when the alternatives a, b are *independent* of a profile $\boldsymbol{P}$, i.e., when for each preference in $\boldsymbol{P}$, a and b might be compared to each other, but neither are compared to any other c (formally, $ac, ca, bc, cb \notin P_i$ for all $c \in A \setminus \{a, b\}, i \in N$). Figure 4.2 shows an example.

$P_1 = \{ab, cd, ce\}$	$P_1 = \{ab, cd, cb\}$
$P_2 = \{de\}$	$P_2 = \{de\}$

Figure 4.2: On the left, a and b are independent of the depicted profile; on the right, they are not (because in P_1 alternative b is also compared to c).

> The reader may wonder why we employ such a strong definition for the independence of two alternatives with respect to a specific profile (which, also, does not directly correspond to the independence definition we used in judgment aggregation). The reason is twofold, bringing together a technical and a conceptual aspect.

Conceptually, we adopt the more stringent notion of independence because it is very intuitive for the purposes of preference aggregation, looking at cases where following the opinion of the majority is undeniably safe and attractive, and leading to a very weak axiom. Weak axioms imply strong characterisation results, showing that a rule is the only one that can be chosen if certain basic conditions are satisfied.

Technically, suppose we replaced the aforementioned independence definition with a weaker one in the spirit of the judgment aggregation model, saying that for a pair ab and a profile $\boldsymbol{P}$, the new profile $\boldsymbol{P}_{\langle ab \rangle}$ in which alternatives a and b are switched consists of acyclic preferences only. With this definition, we would not be able to prove all characterisation theorems of this section regarding constant weights (Theorem 4.12 for aggregation rules that output linear orders would go through, but Theorems 4.13 and 4.14 would cease to hold). Consider, for example, a profile with three agents and three alternatives, $\boldsymbol{P} = (\{ba, ac, bc\}, \{ac\}, \{ac\})$. Although ba is accepted by a majority of agents, the winner-type constant weight rule would select alternative a as the winner, and ba would not be in the collective outcome.

In Section 4.1.1 on judgment aggregation, we discussed three versions of majoritarianism, all of which characterised the constant weight rule under certain assumptions on the superpopulation size. Let us inspect them in the context of preference aggregation.

First, *forward majoritarianism* requires that for every pair of alternatives such that a majority of the agents agrees about their relative ranking, all collective outcomes should respect that ranking. This axiom makes sense as is for collective outcomes that are linear orders, i.e., that rank every two alternatives. Indeed, Theorem 4.12 below is the counterpart of Theorem 4.5, which was proven for judgment aggregation.

Axiom: ***Forward Majoritarianism.*** *For every profile $\boldsymbol{P} \in \mathcal{P}(A)^n$ and alternatives $a, b \in A$ that are independent of $\boldsymbol{P}$, the following holds:*

$$\text{if } ab \in m(\boldsymbol{P}), \text{ then } ab \in P \text{ for all } P \in F(\boldsymbol{P})$$

However, asking for a specific ranked pair to be included in the collective decision is not reasonable when that decision is incomplete (e.g., when it concerns some winning alternatives rather than a full preference). In that case, we slightly modify the axiom, so that it captures the same main idea: We say that for every pair of alternatives such that a majority of the agents agrees about their relative ranking, all collective outcomes should not directly violate that ranking. We call this axiom weak forward majoritarianism.

Axiom: ***Weak Forward Majoritarianism.*** *For every profile $\boldsymbol{P} \in \mathcal{P}(A)^n$ and alternatives $a, b \in A$ that are independent of $\boldsymbol{P}$, the following holds:*

$$\text{if } ab \in m(\boldsymbol{P}), \text{ then } ba \notin P \text{ for all } P \in F(\boldsymbol{P})$$

For aggregation rules that always yield complete outcomes, weak forward majoritarianism coincides with forward majoritarianism.

In judgment aggregation, our characterisation results also required unconstrainedness of the agenda. All technical conditions that were guaranteed to hold by that assumption are now fulfilled for a set of alternatives structured in terms of acyclic preferences, so no additional assumptions are needed.

4.12. Theorem. *For any infinite set of potential alternatives $A^\star$ and for any infinite superpopulation $N^\star$, a weight rule of type $\mathcal{L}$ satisfies forward majoritarianism if and only if it is the constant weight rule $F_c^{\mathcal{L}}$.*

Proof. $F_c^{\mathcal{L}}$ obviously satisfies forward majoritarianism. For the other direction, consider an arbitrary weight function w that induces a weight rule $F_w^{\mathcal{L}}$ and suppose that $F_w^{\mathcal{L}}$ satisfies forward majoritarianism. To prove that $F_w^{\mathcal{L}}$ is the constant weight rule it suffices to show that $w_\lambda = w_{\lambda'}$ for all $\lambda, \lambda' \in \mathbb{N}$.

We first show that $F_w^{\mathcal{L}}$ satisfying forward majoritarianism for a group $N \subseteq N^\star$ with odd $n > 1$ implies that $\frac{w_\lambda}{w_{\lambda'}} < 1 + \frac{2}{n-1}$ for all λ, λ'. Indeed, consider two arbitrary $\lambda > \lambda'$ and a set of alternatives $A = \{a_1, b_1, a'_1, b'_1, \dots, a_\lambda, b_\lambda, a'_\lambda, b'_\lambda\} \subseteq A^\star$ (it is possible to select any suitable number of alternatives because $A^\star$ is assumed to be infinite). Then, take two preference sets as follows:

$$P = \{a_1 b_1, a_2 b_2, \dots, a_\lambda b_\lambda\} \in \mathcal{P}(A)$$
$$P' = \{b_1 a_1, a'_2 b'_2, \dots, a'_{\lambda'} b'_{\lambda'}\} \in \mathcal{P}(A)$$

We construct the following profile $\boldsymbol{P} \in \mathcal{P}(A)^n$ for an odd $n > 1$:

$$\boldsymbol{P} = (\underbrace{P, \dots, P}_{(n-1)/2}, \underbrace{P', \dots, P'}_{(n+1)/2})$$

Clearly, a_1, b_1 are independent of $\boldsymbol{P}$, and $b_1 a_1 \in m(\boldsymbol{P})$. Hence, by forward majoritarianism it must be the case that $b_1 a_1 \in P''$ for all $P'' \in F_w^{\mathcal{L}}(\boldsymbol{P})$, implying that $\frac{n-1}{2} w_{|P|} < \frac{n+1}{2} w_{|P'|}$ and thus $\frac{w_\lambda}{w_{\lambda'}} < 1 + \frac{2}{n-1}$.

To conclude, we know that if $F_w^{\mathcal{L}}$ satisfies forward majoritarianism, then it does so for every group $N \subseteq N^\star$ of odd cardinality. Thus, we must have that $\frac{w_\lambda}{w_{\lambda'}} < 1 + \frac{2}{n-1}$ for every odd $n > 1$ and for every $w_\lambda, w_{\lambda'} \in \mathbb{R}^+$. Letting n go to infinity, this implies that $w_\lambda = w_{\lambda'}$ for all λ, λ'. $\square$

As we have proven in the context of judgment aggregation (Proposition 4.4), having a finite superpopulation ensures that the weights will be very close, but not equal to each other—for convenience, we here only elaborate on the infinite superpopulation case. Interestingly, the characterisation of constant weights under an infinite superpopulation is robust across different types of outcomes.

4.13. Theorem. *For any infinite set of potential alternatives $A^\star$ and for any infinite superpopulation $N^\star$, a weight rule of type $\mathcal{W}$ satisfies weak forward majoritarianism if and only if it is the constant weight rule $F_c^{\mathcal{W}}$.*

Proof. The proof matches that of Theorem 4.12, so we only explain here the part that differs, after the construction of the profile $\boldsymbol{P}$. Let P^a denote the preference set in $\mathcal{W}(A)$ corresponding to a being the winning alternative. We have the following:

$$\sum_{i\in N} w_{P_i}(P^{a_1}) = \sum_{i\in N} w_{P_i}(P^{a_2}) = \cdots = \sum_{i\in N} w_{P_i}(P^{a_\lambda}) = \frac{n-1}{2} w_\lambda$$

and

$$\sum_{i\in N} w_{P_i}(P^{b_1}) = \sum_{i\in N} w_{P_i}(P^{a'_2}) = \cdots = \sum_{i\in N} w_{P_i}(P^{a'_{\lambda'}}) = \frac{n+1}{2} w_{\lambda'}$$

Thus, either each of $a_1, a_2, \ldots, a_\lambda$ win in some outcome of the profile $\boldsymbol{P}$ or none of them do, and the same is true for $b_1, a'_2, \ldots, a'_{\lambda'}$. This means that it cannot be the case that both $a_1b_1 \notin P''$ and $b_1a_1 \notin P''$ for all $P'' \in F_w^{\mathcal{W}}(\boldsymbol{P})$, because then $\emptyset \in F_w^{\mathcal{W}}(\boldsymbol{P})$, which would be contrary to the definition of a weight rule. But a_1, b_1 are independent of $\boldsymbol{P}$, and $b_1a_1 \in m(\boldsymbol{P})$. Weak forward majoritarianism entails that $b_1a_1 \in P''$ for some $P'' \in F_w^{\mathcal{W}}(\boldsymbol{P})$. The remainder of the proof proceeds as for Theorem 4.12. □

4.14. Theorem. *For any infinite set of potential alternatives $A^\star$ and for any infinite superpopulation N, a weight rule of type $\mathcal{W}_k$ satisfies weak forward majoritarianism if and only if it is the constant weight rule $F_c^{\mathcal{W}_k}$.*

Proof. Analogous to the proof of Theorem 4.13 above. □

After concluding our study of forward majoritarianism for preferences aggregation settings, let us proceed to the other two axioms, encoding different aspects if the idea of respecting majority opinions, which we met in the previous section on judgment aggregation, namely, backward majoritarianism and general majoritarianism. Loosely speaking, backward majoritarianism prescribes that when an ordered paid of alternatives ab appears in the collective outcome, this should be because a majority of the agents have ranked a above b. Except for outcomes that are linear orders (in which case the results and proofs go through similarly as for the judgment aggregation model), this property does not make much sense in the context of voting, and it definitely does not characterise the constant weight rule, as Example 4.15 illustrates.

4.15. Example. Consider alternatives a, b, and c, a profile of two agents with preferences $\{ac\}$ and $\{ac\}$, respectively, and the constant weight rule of single-winner type. Then, alternative a will be the winner, and the pair ab will be included in the outcome, despite the fact that no majority of agents have ab in their preference set. △

The axiom of general majoritarianism is even less appropriate for single-winner and multiple-winner outcomes, which are incomplete by definition, as it presumes completeness of the collective opinions.

Splitting

Recall the central motivation behind the splitting axioms in judgment aggregation: the agents should not have the possibility to alter the final outcome by submitting a larger consistent judgment, induced by the combination of their individual judgments. This idea extends to preference aggregation as well, supposing that the agents could report the union of the individual preference sets, when the corresponding new preference remains acyclic. This is useful in practice: If some members of a council meet before a vote and realise that they do not disagree on any of the relative rankings between alternatives, they could send a messenger to the main meeting to vote on their behalf.

We now state the splitting axioms for preference aggregation. All axioms and relevant results hold for the three types of weight rules in which we are interested, i.e., for outcomes that are linear orders, or that denote a (set of) winner(s).

Axiom: ***Splitting, Equal Splitting, and Single Splitting.*** *For every profile $\boldsymbol{P} \in \mathcal{P}(A)^n$ and subgroup $N' \subseteq N$ of agents with pairwise disjoint ballots, it is the case that $\bigcup_{i\in N'} P_i \in \mathcal{P}(A)$ implies $F(\boldsymbol{P}) = F(\boldsymbol{P}')$, where $\boldsymbol{P}'$ arises from $\boldsymbol{P}$ by replacing the ballot of each member of N' by the union $\bigcup_{i\in N'} P_i$. F furthermore is said to satisfy equal splitting if the ballots of all agents in N' are of equal size, and single splitting if they are all singleton sets.*

4.16. Theorem. *For any infinite set of potential alternatives $A^\star$, any infinite superpopulation $N^\star$, and any type $\mathcal{T} \in \{\mathcal{L}, \mathcal{W}, \mathcal{W}_k\}$, a weight rule of type $\mathcal{T}$ satisfies single splitting if and only if it is the equal-and-even weight rule $F^{\mathcal{T}}_{ee}$.*

Proof. Given a type $\mathcal{T} \in \{\mathcal{L}, \mathcal{W}, \mathcal{W}_k\}$, we consider an arbitrary instance of the single splitting axiom and show that $F^{\mathcal{T}}_{ee}$ satisfies it. Take two profiles $\boldsymbol{P}, \boldsymbol{P}'$, where $\boldsymbol{P}'$ arises from $\boldsymbol{P}$ by replacing the preference set of each member of N' by the union $\bigcup_{i\in N'} P_i$. By the definition of the weights for the rule $F^{\mathcal{T}}_{ee}$, we know that every pair of alternatives in $\bigcup_{i\in N'} P_i$ weighs exactly the same in $\boldsymbol{P}$ and in $\boldsymbol{P}'$, because $|P_i| = 1$ for every $i \in N'$, and $w_\lambda = \frac{w_1}{\lambda}$ for every λ. The same holds for all pairs that are not in $\bigcup_{i\in N'} P_i$ too, since they trivially appear in exactly the same preference sets in both profiles $\boldsymbol{P}$ and $\boldsymbol{P}'$. This means that $F^{\mathcal{T}}_{ee}(\boldsymbol{P}) = F^{\mathcal{T}}_{ee}(\boldsymbol{P}')$.

For the other direction, consider an arbitrary $\lambda \in \mathbb{N}$, a set of alternatives $A = \{a, b, c_1, c_2, \ldots, c_{\lambda-2}, c_{\lambda-1}\} \subseteq A^\star$ (which is possible because $A^\star$ is infinite), and two profiles $\boldsymbol{P}$ and $\boldsymbol{P}'$ as follows, for a sufficiently large group $N \subseteq N^\star$:

$$\boldsymbol{P} = (\{ab\}, \{ba\}, \{c_1b\}, \{c_2b\} \ldots, \{c_{\lambda-2}b\}, \{c_{\lambda-1}b\})$$
$$\boldsymbol{P}' = (\{ab\}, \underbrace{\{ba, c_1b, c_2b \ldots, c_{\lambda-2}b, c_{\lambda-1}b\}, \ldots, \{ba, c_1b, c_2b \ldots, c_{\lambda-2}b, c_{\lambda-1}b\}}_{\lambda})$$

First, for any weight rule $F^{\mathcal{L}}_w$ we must have that $L, L' \in F^{\mathcal{L}}_w(\boldsymbol{P})$ for some preferences $\{ab, c_1b, \ldots, c_{\lambda-2}b, c_{\lambda-1}b\} \subseteq L$ and $\{ba, c_1b, \ldots, c_{\lambda-2}b, c_{\lambda-1}b\} \subseteq L'$. Moreover, if

$F_w^{\mathcal{L}}$ satisfies single splitting, it must be the case that $F_w^{\mathcal{L}}(\boldsymbol{P}) = F_w^{\mathcal{L}}(\boldsymbol{P}')$. So the weights of ab and ba should be the same in the profile $\boldsymbol{P}'$, that is, $w_1 = \lambda \cdot w_\lambda$.

Second, for any weight rule $F_w^{\mathcal{W}}$, the preference orders that have some alternative from the set $\{a, b, c_1, c_2, \ldots, c_{\lambda-2}, c_{\lambda-1}\}$ as the winner must all be in $F_w^{\mathcal{W}}(\boldsymbol{P})$. By the satisfaction of single splitting, we should have that $F_w^{\mathcal{L}}(\boldsymbol{P}) = F_w^{\mathcal{L}}(\boldsymbol{P}')$, and hence the weights of ab and ba should be the same in the profile $\boldsymbol{P}'$. So, $w_1 = \lambda \cdot w_\lambda$. The same argument holds for a weight rule $F_w^{\mathcal{W}_k}$ too. □

4.17. Theorem. *For any infinite set of potential alternatives $A^\star$, any infinite superpopulation $N^\star$, and any type $\mathcal{T} \in \{\mathcal{L}, \mathcal{W}, \mathcal{W}_k\}$, a weight rule of type $\mathcal{T}$ satisfies equal splitting if and only if it is the equal-and-even weight rule $F_{ee}^{\mathcal{T}}$.*

Proof. If a weight rule satisfies equal splitting, then it satisfies single splitting too—by Theorem 4.7, it must be the equal-and-even-scoring rule. The proof of the other direction proceeds as the first part of the proof of Theorem 4.7. □

4.18. Proposition. *For any set of potential alternatives $A^\star$ and any superpopulation $N^\star$ such that $|N^\star| > 2$ and $|A^\star| > 5$, no weight rule of type $\mathcal{T}$ satisfies arbitrary splitting, for $\mathcal{T} \in \{\mathcal{L}, \mathcal{W}, \mathcal{W}_k\}$.*

Proof. Every weight rule of type $\mathcal{T}$ that satisfies arbitrary splitting also satisfies single splitting, so by Theorem 4.16, the only candidate rule is $F_{ee}^{\mathcal{L}}$. Then it remains to show that $F_{ee}^{\mathcal{T}}$ fails arbitrary splitting on at least one counterexample. The counterexample is similar to the one we presented in the proof of Proposition 4.9, translated from judgment aggregation to preference aggregation.

Let $A = \{a_1, b_1, a_2, b_2, a_3, b_3\}$ and consider two profiles:

$$\boldsymbol{P} = (\{b_1a_1\}, \{a_1b_1\}, \{a_2b_2, a_3b_3\})$$
$$\boldsymbol{P}' = (\{b_1a_1\}, \{a_1b_1, a_2b_2, a_3b_3\}, \{a_1b_1, a_2b_2, a_3b_3\})$$

Let w be the weight function associated with $F_{ee}^{\mathcal{T}}$. Since both a_1b_1 and b_1a_1 weigh w_1 in profile $\boldsymbol{P}$, they will both belong to the collective outcome (either that outcome taking the form of a linear order, or denoting single or multiple winners). Note that $w_{|\{a_1b_1, a_2b_2, a_3b_3\}|} = \frac{w_1}{3}$. So, a_1b_1 weighs $\frac{2w_1}{3} < w_1$ in $\boldsymbol{P}'$, while b_1a_1 still weighs w_1. This means that for every preference set in $F_{ee}^{\mathcal{T}}(\boldsymbol{P}')$, only b_1a_1 but not a_1b_1 will belong to that set. Hence, $F_{ee}^{\mathcal{L}}(\boldsymbol{P}) \neq F_{ee}^{\mathcal{L}}(\boldsymbol{P}')$, which violates the arbitrary splitting axiom. □

Overall, majoritarianism and splitting can never be simultaneously satisfied, by any weight rule for preference aggregation.

4.19. Corollary. *For any infinite set of potential alternatives, any infinite superpopulation $N^\star$, and any type $\mathcal{T} \in \{\mathcal{L}, \mathcal{W}, \mathcal{W}_k\}$, there is no weight rule $F_w^{\mathcal{T}}$ that satisfies weak forward majoritarianism together with some form of splitting.*

Finally, note that during the axiomatic study of weight rules in the setting of judgment aggregation (in Section 4.1.1), we also examined the property of *quality over quantity*, requiring that the collective outcome follow the opinion of the agents with smaller sets of judgments. To prove that the described property characterises a rule that we call *upward lexicographic*, we made use of a finite superpopulation. In the current section on preference aggregation, all our formal results call for the assumption of an infinite superpopulation, and it would not be appealing to suddenly alter it. But more importantly, the axiom of quality over quantity does not fit in our preference aggregation model that relies on types: it would be unreasonable to demand of a rule that outputs preferences that are of winner type to directly copy the preference of an agent that can take any arbitrary form. We thus do not involve this axiom in our analysis of weight rules for voting and preference aggregation.

Figure 4.3 below summarises the main message of this section, regarding axiomatic properties for weight rules, abstracting away from all technical details.

Desirable axiom		Suitable rule
majoritarianism	$\leftrightarrow$	constant weight
splitting	$\leftrightarrow$	equal-and-even weight

Figure 4.3: High-level correspondence between axioms and weight rules.

4.2 On Quota Rules (Judgment Aggregation)

Turning again to the aggregation of incomplete judgments, in this section we will examine the class of quota rules that we defined in Section 3.2, from an axiomatic point of view. We will not only describe the properties that single out quota rules amongst all judgment aggregation rules, but we will also identify those axioms that characterise different kinds of rules within the general class of quota rules.

To warm up, let us revisit an axiom that was central in the previous section on weight rules, namely the splitting axiom, but now in the context of quota rules:

4.20. Example. Consider the two profiles $\boldsymbol{J}$ (left) and $\boldsymbol{J}'$ (right) depicted below.

The individual judgments of our three agents in profile $\boldsymbol{J}$ are clearly disjoint and logically consistent, so they could be merged and lead to the profile $\boldsymbol{J}'$. Suppose we use the strict majority rule: A quota rule that associates an invariable absolute threshold of 2 with each one of the propositions p, q, and $p \wedge q$ (that is, each proposition will be collectively accepted if only if at least two out of the three agents include it in their individual judgments). In profile $\boldsymbol{J}$, the rule will only take into account one agent supporting each proposition, and will not accept it; but in profile $\boldsymbol{J}'$, the rule will

	p	q	$p \wedge q$
Ann:	☑	☐	☐
Bea:	☐	☑	☐
Cal:	☐	☐	☑
Majority:	☐	☐	☐

	p	q	$p \wedge q$
Ann:	☑	☑	☑
Bea:	☑	☑	☑
Cal:	☑	☑	☑
Majority:	☑	☑	☑

see three agents supporting each proposition and the threshold of acceptance will be reached. Hence, splitting will be violated.

In general, the splitting axiom will be violated by all quota rules, unless we have a trivial such rule where every proposition is associated either with the threshold 0 or with the threshold $n+1$ (meaning that it will always, or never, be accepted, respectively). $\triangle$

In this section, we will employ axioms that we have not met in this thesis yet, but that nonetheless are common in the social choice literature. The characterisation theorems that we will present are in line with known results that have been proven for agents who all report complete judgments (Dietrich and List, 2007b). In particular, we also address the obstacles that may emerge when a given quota rule produces inconsistent outcomes (recall the discussion we had in the context of Example 3.17, page 34). The well-known trade-off in judgment aggregation between independent (issue-by-issue) aggregation and logical consistency evidently manifests itself in quota rules. Dietrich and List (2010a) have also studied a special quota rule, namely the majority rule, for the general case of possibly incomplete individual judgments and regarding the problem of logical inconsistency. Specifically, Dietrich and List have explored what domains of decision making (i.e., agendas) are suitable for consistent majority aggregation—in this section, we do not impose any constraints on the domain, but instead study what quota rules guarantee consistent outcomes given the structure of any existing agenda.

Leaving aside the logical interactions between the issues at stake, there is a large pool of axiomatic literature involved with referendums, and predominantly with votes on single binary choices. This stream of work was pioneered by May (1952), who proved that the majority rule is the only rule simultaneously satisfying the axioms of *anonymity* (i.e., treating all agents symmetrically), *neutrality* (i.e., treating all propositions symmetrically), and *monotonicity* (i.e., increasing the support for a proposition is not harmful for the collective decision on that proposition). For more than one issue, anonymity and monotonicity are also central in the axiomatisations of quota rules, while neutrality would force the thresholds of acceptance for the different propositions to coincide. Note that this first work on the topic by May hinged on complete individual opinions, but more recently scholars have relaxed this assumption.

Characteristically, Côrte-Real and Pereira (2004) have analysed systems of referendums used by countries in the European Union by employing an axiomatic methodology. They were specifically interested in a version of the non-show paradox (that is, a situation where an agent can improve the collective decision for herself by abstaining), which will also play a role later on in this thesis, in Section 5.1 of Chapter 5. Many of their results can be translated in our model for the special case of a single issue.

Other axiomatic works related to this section have been conducted by Llamazares (2006) and by Houy (2007). Both authors have examined voting rules on a single binary choice, defined with respect to the difference in number between those that support the given choice and those that oppose to it—they call this kind of aggregation methods *majority of difference*.[4] In our framework, such rules are called *marginal quota rules* in order to stay as close as possible to the—established in judgment aggregation—term *quota rules*. Considering the special case of judgment aggregation on a single proposition, the our characetrisation results of this section generalise the characterisations obtained by Llamazares and Houy. Then, variants of the independence axiom in judgment aggregation essentially guarantee that in the most general case with several interconnected issues, quota rules apply in an issue-by-issue basis. Overall, this section builds important links between two previously separate domains of research: first, single-issue voting rules with majority quota, and second, quota rules in judgment aggregation, for multiple interrelated issues.

> For the whole section, we fix an agenda $\Phi \subseteq \Phi^\star$ and a group of agents $N \subseteq N^\star$ (of size n) in order to define the properties of an aggregation rule. More precisely, we want the relevant properties to hold for all such agendas and groups of agents. Note also that our results of this section do not depend on specific assumptions about the superagenda $\Phi^\star$ and the superpopulation $N^\star$, which is why we omit such references in the statements of our theorems.

4.2.1 Axioms

As already mentioned, anonymity is a classical axiom, initially defined for complete individual judgments, that can be naturally defined for the incomplete case.

Axiom: ***Anonymity (A).*** *For all permutations $\pi : N \to N$ and profiles $\boldsymbol{J} = (J_1, \dots, J_n)$, it holds that $F(\boldsymbol{J}) = F(\pi(\boldsymbol{J}))$, where $\pi(\boldsymbol{J}) = (J_{\pi(1)}, \dots, J_{\pi(n)})$.*

We next explore how other desirable properties of aggregation rules may be defined, extending their axiomatic counterparts for the complete case by specifically taking into account the potential incompleteness of the judgments.

We begin with the property of *monotonicity*, which, broadly speaking, states that extra support on a given proposition φ should never harm the collective acceptance of that proposition. Two clarifications are in order here. First, what exactly does the term "extra support" mean? Second, when is the collective acceptance of a proposition considered "harmed"? Relevant to the first question is the observation that an agents can accept a proposition φ that she was previously rejecting in two scenarios: (i) by

[4] Such binary voting rules were already mentioned by Fishburn (1973).

including φ in her judgment that was before abstaining on the issue $\widetilde{\varphi}$ and (ii) by including φ in her judgment that was before containing the proposition $\neg\varphi$. The former action could be seen as less radical than the latter. Moreover, an agent can possibly promote proposition φ even more indirectly, by abstaining on $\widetilde{\varphi}$ instead of accepting $\neg\varphi$. Similarly, and regarding the second question above, a proposition φ can disappear from the collective outcome because $\neg\varphi$ is collectively accepted instead, or also independently of that (since collective judgment sets are not necessarily consistent, it may be the case that both φ and $\neg\varphi$ are included in the outcome). The following definition will help us make these ideas concrete:

4.21. Definition. *Given two judgment sets J and J' and a proposition φ, the support of φ **weakly increases** from J to J' (and weakly decreases from J' to J), denoted by $J \leqslant_\varphi J'$, if $J \cap \{\varphi\} \subseteq J' \cap \{\varphi\}$. The relation $<_\varphi$ is defined as the strict part of $\leqslant_\varphi$. We write $J \doteq_\varphi J'$ if and only if $J \leqslant_\varphi J'$ and $J' \leqslant_\varphi J$.*

It is implied that $J' \leqslant_\varphi J$ if and only if it is not the case that $J <_\varphi J'$.

Numerous versions of the monotonicity property can be defined, depending on the precise way we interpret the notion of support of a proposition. Here we discuss two such versions, a strong one and a weak one.[5]

Axiom: ***Judgment Monotonicity (M).*** *For all propositions $\varphi \in \Phi$ and all profiles $\boldsymbol{J} = (J_1, \ldots, J_i, \ldots, J_n)$ and $\boldsymbol{J'} = (J_1, \ldots, J'_i, \ldots, J_n)$, the following holds:*

$$J_i <_\varphi J'_i \quad \textit{implies that} \quad F(\boldsymbol{J}) \leqslant_\varphi F(\boldsymbol{J'}) \textit{ and } F(\boldsymbol{J'}) \leqslant_{\neg\varphi} F(\boldsymbol{J})$$

Axiom: ***Weak Judgment Monotonicity (WM).*** *For all propositions $\varphi \in \Phi$ and all profiles $\boldsymbol{J} = (J_1, \ldots, J_i, \ldots, J_n)$ and $\boldsymbol{J'} = (J_1, \ldots, J'_i, \ldots, J_n)$, the following holds:*

$$\neg\varphi \in J_i \textit{ and } \varphi \in J'_i \quad \textit{imply that} \quad F(\boldsymbol{J}) \leqslant_\varphi F(\boldsymbol{J'})$$

When the input profiles are complete, judgment monotonicity and weak judgment monotonicity reduce to the same axiom, viz., the standard monotonicity axiom in judgment aggregation (Dietrich and List, 2007b).

We continue with *independence*, an axiom requiring that the positive judgments on $\widetilde{\varphi}$ and only those play a role in the collective decision about φ. The importance of independence becomes evident for agendas with more than one issue, and will play a determining role in our analysis of quota rules—indeed quota rules are among the few natural judgment aggregation rules that satisfy independence in the complete case. With respect to the incomplete case, recall the weight rules that we defined in Section 3.1.1. Weight rules select a collective judgment by maximising the total weight contributed by the agents, considering all propositions simultaneously, an idea that stands in opposition to that of independence.

The basic concept of independence has two implicit parts:

[5]With a single-issue agenda as a departing point, weak judgment monotonicity corresponds to the axiom that Houy (2007) calls "weak monotonicity 2".

(i) Judgments on $\neg\varphi$ should not affect the collective outcome about φ;

(ii) Judgments on $\widetilde{\psi} \neq \widetilde{\varphi}$ should not affect the collective outcome about φ.

When individual judgments are complete, case (ii) is the only interesting one. But if we allow for incompleteness, this is not true anymore. We thus define, additionally to the most general independence property, a weaker version.[6]

Axiom: ***Independence (I).*** *For all propositions $\varphi \in \Phi$ and for all profiles $\boldsymbol{J}, \boldsymbol{J}'$, the following holds:*

$$N_\varphi^{\boldsymbol{J}} = N_\varphi^{\boldsymbol{J}'} \quad \textit{implies that} \quad F(\boldsymbol{J}) \doteq_\varphi F(\boldsymbol{J}')$$

Axiom: ***Weak Independence (WI).*** *For all propositions $\varphi \in \Phi$ and for all profiles $\boldsymbol{J}, \boldsymbol{J}'$, the following holds:*

$$N_\varphi^{\boldsymbol{J}} = N_\varphi^{\boldsymbol{J}'} \textit{ and } N_{\neg\varphi}^{\boldsymbol{J}} = N_{\neg\varphi}^{\boldsymbol{J}'} \quad \textit{imply that} \quad F(\boldsymbol{J}) \doteq_\varphi F(\boldsymbol{J}')$$

Notably, the two variants of monotonicity coincide under independence.

4.22. Proposition. *Judgment monotonicity is logically equivalent to weak judgment monotonicity together with independence.*

Proof. From the relevant definitions it follows that judgment monotonicity implies weak judgment monotonicity. Now, we will also show that weak judgment monotonicity together with independence implies judgment monotonicity. Consider two profiles:

$$\boldsymbol{J} = (J_1, \ldots, J_i, \ldots, J_n) \qquad \text{and} \qquad \boldsymbol{J}' = (J_1, \ldots, J_i', \ldots, J_n)$$

and suppose that $J_i <_\varphi J_i'$. For judgment monotonicity to hold, we need to prove that $F(\boldsymbol{J}) \leq_\varphi F(\boldsymbol{J}')$ and that $F(\boldsymbol{J}') \leq_{\neg\varphi} F(\boldsymbol{J})$. We will show the former (and the latter can be shown symmetrically).

If $\neg\varphi \in J_i$ and $\varphi \in J_i'$, we are done by weak judgment monotonicity.

If $\varphi, \neg\varphi \notin J_i$ and $\varphi \in J_i'$, we create a new profile based on $\boldsymbol{J}$ by only changing the judgment of agent i as follows:

$$\boldsymbol{J}'' = (J_1, \ldots, \{\neg\varphi\}, \ldots, J_n)$$

By independence $F(\boldsymbol{J}) \doteq_\varphi F(\boldsymbol{J}'')$, and by weak judgment monotonicity $F(\boldsymbol{J}'') \leq_\varphi F(\boldsymbol{J}')$. Thus, we have that $F(\boldsymbol{J}) \leq_\varphi F(\boldsymbol{J}')$. □

Next, the axiom of *judgment cancellation* is particularly relevant when modelling incomplete individual judgments, and suggests that adding to a profile the same number of agents supporting a proposition φ and its negation $\neg\varphi$ should not change the collective outcome on that proposition. Llamazares (2006) also defines cancellation within

[6] Weak independence has previously been defined by Gärdenfors (2006) and Dietrich and List (2008).

binary voting with one issue. However, the setting of this section is more engaged, since the multiplicity of issues brings to light original interactions between the axioms of cancellation and independence.[7]

AXIOM: ***Judgment Cancellation (C).*** *For all propositions* $\varphi \in \Phi$ *and for all profiles* $\boldsymbol{J} = (J_1, \ldots, J_i, \ldots, J_j, \ldots, J_n)$ *and* $\boldsymbol{J}' = (J_1, \ldots, J'_i, \ldots, J'_j, \ldots, J_n)$,

$$\varphi, \neg\varphi \notin J_i \cup J_j \textit{ and } \varphi \in J'_i, \neg\varphi \in J'_j \textit{ imply that } F(\boldsymbol{J}) \doteq_\varphi F(\boldsymbol{J}')$$

> Both judgment monotonicity and judgment cancellation bear a flavour of weak independence: They stress that we can restrict attention to a relevant proposition φ and inspect two profiles where some of the submitted judgments on different propositions ψ may change. This is a debatable feature of these definitions. Nonetheless, we insist on using them in order to facilitate the comparison of this work with two central results of the judgment aggregation literature where monotonicity is defined in an analogous manner: the axiomatisation result of quota rules for complete inputs (Dietrich and List, 2007b), and the characterisation of all rules immune to manipulation (Dietrich and List, 2007c). Besides keeping the connections with previous work, note that the definitions we use do not have substantial technical impact on the results of this section, since all quota rules are already weakly independent.

Finally, quota rules in general do not satisfy another popular axiom in judgment aggregation, namely *neutrality* (Terzopoulou and Endriss, 2020). Neutrality demands that all propositions be treated equally by an aggregation rule, while this is not the case for quota rules that assign different thresholds to different propositions. Similarly, quota rules violate *unbiasedness* (sometimes called acceptance-rejection neutrality), which imposes the equal treatment between a proposition and its negation (Botan et al., 2016).

4.2.2 Characterisation Results

Given the axioms presented in Section 4.2.1, we characterise all different quota rules for incomplete individual judgments. Our proofs are inspired by the proof of Dietrich and List (2007b) for the characterisation of quota rules in the complete framework, which uses the axioms of anonymity, monotonicity, and independence.

4.23. THEOREM. *An aggregation rule for incomplete judgments is an invariable absolute q. rule if and only if it satisfies anonymity, judg. monotonicity, and independence.*

PROOF. This proof directly follows the proof of Theorem 1 by Dietrich and List (2007b), for quota rules in the complete setting. □

[7] For details consult Table 4.1, showing that independence and judgment cancellation are not simultaneously satisfied by any of the quota rules considered here.

4.24. Theorem. *An aggregation rule F for incomplete judgments is an invariable marginal quota rule if and only if it satisfies anonymity, judgment monotonicity, weak independence, and judgment cancellation.*

Proof. To check that invariable marginal quota rules satisfy all the axioms of anonymity, weak independence, judgment monotonicity, and judgment cancellation is easy. For the other direction, consider an arbitrary aggregation rule F satisfying all the axioms of the hypothesis, and an arbitrary proposition $\varphi \in \Phi$. If φ never belongs to the collective outcome, no matter what the input profile is, then we can take $m_\varphi = n+1$. Otherwise, there exists a profile $\boldsymbol{J}$ such that $\varphi \in F(\boldsymbol{J})$. So, we can consider a specific such profile $\boldsymbol{J} \in \operatorname{argmin}_{\boldsymbol{J}:\varphi\in F(\boldsymbol{J})} n_\varphi^{\boldsymbol{J}} - n_{\neg\varphi}^{\boldsymbol{J}}$ and define the number m_φ as follows:

$$m_\varphi = \min_{\boldsymbol{J}|\varphi\in F(\boldsymbol{J})} n_\varphi^{\boldsymbol{J}} - n_{\neg\varphi}^{\boldsymbol{J}}$$

From anonymity and weak independence, it follows that φ will belong to the collective outcome for every profile where propositions φ and $\neg\varphi$ have the same number of supporters as in $\boldsymbol{J}$. Then, judgment monotonicity implies that φ will belong to the collective outcome also for every profile in which the support of φ increases or the support of $\neg\varphi$ decreases with respect to the relevant support in $\boldsymbol{J}$. Moreover, the axiom of judgment cancellation suggests that by increasing (or decreasing) the support of φ and $\neg\varphi$ to the same degree, φ will still belong to the collective outcome. Formally, we have that $\varphi \in F(\boldsymbol{J}')$ for all profiles $\boldsymbol{J}'$ with $n_\varphi^{\boldsymbol{J}'} - n_{\neg\varphi}^{\boldsymbol{J}'} \geqslant n_\varphi^{\boldsymbol{J}} - n_{\neg\varphi}^{\boldsymbol{J}} = m_\varphi$. Finally, by definition of m_φ, we have that $\varphi \notin F(\boldsymbol{J}')$ for all profiles $\boldsymbol{J}''$ with $n_\varphi^{\boldsymbol{J}''} - n_{\neg\varphi}^{\boldsymbol{J}''} < n_\varphi^{\boldsymbol{J}} - n_{\neg\varphi}^{\boldsymbol{J}} = m_\varphi$. So, F coincides on φ with the invariable marginal quota rule associating with thresholds m_φ. $\square$

4.25. Theorem. *An aggregation rule for incomplete judgments is a variable quota rule if and only if it satisfies anonymity, weak judg. monotonicity, and weak independence.*

Proof. We will prove the statement for variable absolute quota rules. To check that variable absolute quota rules satisfy all the axioms of anonymity, weak independence, and weak judgment monotonicity is easy. For the other direction, we will consider an arbitrary rule F and we will repeat the following argument for all numbers $k \in \{0,\dots,n\}$: If φ never belongs to the collective outcome for profiles $\boldsymbol{J}$ with $n_{\tilde{\varphi}}^{\boldsymbol{J}} = k$, then we can take $m_\varphi^k = k+1$. Otherwise, there exists a profile $\boldsymbol{J}$ where exactly k agents report an opinion on φ such that $\varphi \in F(\boldsymbol{J})$. We consider a specific such profile $\boldsymbol{J}$ with the smallest number of supporters of φ and define the number m_φ^k as follows:

$$m_\varphi^k = \min_{\substack{\boldsymbol{J}|\varphi\in F(\boldsymbol{J})\\ n_{\tilde{\varphi}}^{\boldsymbol{J}}=k}} n_\varphi^{\boldsymbol{J}}$$

From anonymity and weak independence, we have that φ will belong to the collective outcome for every profile where proposition φ has the same number of supporters as in $\boldsymbol{J}$ and the number of agents reporting a judgment on the issue $\widetilde{\varphi}$ remains the same. Then, weak judgment monotonicity implies that φ will belong to the collective outcome also for every profile in which the support for φ increases while still the number of agents reporting a judgment on the issue $\widetilde{\varphi}$ remains the same. Formally, $\varphi \in F(\boldsymbol{J}')$ for all profiles $\boldsymbol{J}'$ with $n_{\varphi}^{\boldsymbol{J}'} \geqslant n_{\varphi}^{\boldsymbol{J}} = m_{\varphi}^{k}$ and $n_{\widetilde{\varphi}}^{\boldsymbol{J}} = k$. Finally, by definition of m_{φ}, we have that $\varphi \notin F(\boldsymbol{J}')$ for all profiles $\boldsymbol{J}'$ with $n_{\varphi}^{\boldsymbol{J}'} < m_{\varphi}^{k}$ and $n_{\widetilde{\varphi}}^{\boldsymbol{J}} = k$. Hence, F coincides on φ with the variable absolute quota rule associating with φ (in profiles with k agents reporting a judgment on $\widetilde{\varphi}$) the threshold m_{φ}^{k}. □

Freixas and Zwicker (2009) have obtained an analogous result to Theorem 4.23, using anonymity and monotonicity to characterise a class of quota rules that always return a *yes*/*no* answer in single-issue voting. Theorems 4.24 and 4.25 generalise the results of Llamazares (2006) and Houy (2007), respectively, which—restricted to the case of single-issue voting—make use of the same axioms as the ones appearing in the above results, except for independence.[8] Essentially, the original aspect of our characterisations lies in the use of the independence axiom (or versions thereof), which is vacuous in the scope of an agenda with a single issue but prominent in judgment aggregation. Table 4.1 demonstrates succinctly which axioms are satisfied by what rules.

Quota rules \ Axioms	A	M	WM	I	WI	C
invariable absolute	✓	✓	✓	✓	✓	✗
invariable marginal	✓	✓	✓	✗	✓	✓
variable	✓	✗	✓	✗	✓	✗

Table 4.1: Axioms satisfied by quota rules. The coloured cells in each row illustrate the axioms that characterise the relevant class of rules.

> In the characterisation of invariable absolute quota rules we can replace monotonicity with weak monotonicity, since these two versions of judgment monotonicity coincide under independence. On the contrary, weak judgment monotonicity (together with anonymity, weak independence, and judgment cancellation) is not enough to characterise invariable marginal quota rules. Importantly, weak independence cannot replace independence in the characterisation of invariable absolute quota rules.

Let us now see why certain axioms are not satisfied within specific classes of quota rules (as presented in Table 4.1), via a number of examples.

[8]Specifically, Llamazares and Houy add a *Pareto* condition in order to characterise each particular quota rule within the larger classes. Also, Houy incorporates an additional monotonicity axiom that characterises the rules based on quorums among all variable quota rules.

4.26. Example. Consider an invariable absolute threshold $a_\varphi = n-1$ and take a profile where exactly $n-1$ agents accept φ and one agent accepts $\neg\varphi$. Then, φ will belong to the outcome. But it is clear that if the support of both φ and $\neg\varphi$ decreases, then φ will not be in the collective outcome. Thus, judgment cancellation is not always satisfied by invariable absolute quota rules. $\triangle$

4.27. Example. Consider an invariable marginal threshold $m_\varphi = 2$ and take a profile with two agents accepting φ and no agents accepting $\neg\varphi$. In this profile, φ will belong to the collective outcome. But if the support of $\neg\varphi$ increases while that of φ remains the same, then the relevant margin gets smaller and φ will not be in the outcome anymore. Because judgments on $\neg\varphi$ affect proposition φ, independence is not always satisfied by invariable marginal quota rules. $\triangle$

Since the class of variable quota rules includes all the invariable ones, it already follows from the previous examples that judgment cancellation and independence cannot be satisfied by all variable quota rules. In addition, judgment monotonicity can be violated by a variable quota rule as discussed in Example 4.28.

4.28. Example. If $a^2_\varphi = 1$ (that is, when two agents report a judgment on $\widetilde{\varphi}$, one acceptance of φ suffices to have it collectively accepted) and $a^3_\varphi = 3$ (which means that a unanimity is required for φ to be collectively accepted when three agents express a judgment on $\widetilde{\varphi}$), then in a profile $\boldsymbol{J}$ where $n^{\boldsymbol{J}}_\varphi = 1$ and $n^{\boldsymbol{J}}_{\widetilde{\varphi}} = 2$, proposition φ will be in the outcome. But if the support of φ increases by just one agent and we have a new profile $\boldsymbol{J}'$ with $n^{\boldsymbol{J}'}_{\widetilde{\varphi}} = 3$, then proposition φ will not be in the outcome. In such a case, additional support can be damaging to the acceptance of a proposition. $\triangle$

4.2.3 Axioms within Classes of Rules

When we say that an axiom is violated within a particular class of quota rules, what we really mean is that there is *some* rule in this class that provides a counterexample. Next, we answer a more fine-grained question:

> Under what assumptions on the relevant thresholds is an axiom satisfied within a class of rules that generally violate that axiom?

Proposition 4.29 states that no invariable absolute quota rule can satisfy judgment cancellation, unless it is trivial. Put differently, judgment cancellation is the property that differentiates between invariable marginal and invariable absolute quota rules. Analogously (but on the other direction), independence differentiates between invariable absolute and invariable marginal quota rules.

4.29. Proposition. *An invariable absolute (marginal) quota rule satisfies judgment cancellation (independence) if and only if it is trivial.*

Proof. Every trivial quota rule satisfies judgment cancellation. Consider a non-trivial invariable absolute quota rule F_a that induces thresholds a_φ. We know that there exists some $\varphi \in \Phi$ such that $1 \leqslant a_\varphi \leqslant n-1$. Take the profiles $\boldsymbol{J}$ and $\boldsymbol{J}'$ as follows:

$$\boldsymbol{J}: \underbrace{\varphi \ldots \varphi\varphi}_{a_\varphi} \neg\varphi \quad \underbrace{- \ldots -}_{n-a_\varphi-1} \qquad \text{and} \qquad \boldsymbol{J}': \underbrace{\varphi \ldots \varphi}_{a_\varphi-1} - - \quad \underbrace{- \ldots -}_{n-a_\varphi-1}$$

Proposition φ is accepted in $\boldsymbol{J}$ and rejected in $\boldsymbol{J}'$, violating judgment cancellation.

Analogously, only trivial (invariable marginal) quota rules satisfy independence. □

Although the conditions that we will state in Lemma 4.31 are very technical, they help us better understand not only the formal properties of the different quota rules, but also the relations between these rules. Specifically, from (a) and (b) we can deduce that a variable absolute quota rule that satisfies independence will also satisfy judgment monotonicity, and will thus be an invariable absolute quota rule, by Theorem 4.23. So, we obtain Proposition 4.30.[9]

4.30. Proposition. *A variable absolute (or marginal) quota rule satisfies independence if and only if it is an invariable absolute (or marginal) quota rule.*

4.31. Lemma. *A variable absolute quota rule F_{va} associated with thresholds with thresholds a_φ^k satisfies*

(a) *judgment monotonicity if and only if for all $\varphi \in \Phi$ and $k \in \{0, \ldots, n\}$,*

$$\begin{array}{llll} (1) & a_\varphi^{k+1} \leqslant a_\varphi^k + 1 & \textit{when} \quad a_\varphi^k \leqslant k \leqslant n-1 & \textit{and} \\ (2) & a_\varphi^k \leqslant a_\varphi^{k+1} & \textit{when} \quad a_\varphi^{k+1} \leqslant k; & \end{array}$$

(b) *independence if and only if for all $\varphi \in \Phi$ and $k, \ell \in \{0, \ldots, n\}$ with $k < \ell$,*

$$\begin{array}{llll} (1) & a_\varphi^\ell = a_\varphi^k & \textit{when} \quad a_\varphi^k \leqslant k \leqslant n-1 & \textit{and} \\ (2) & k < a_\varphi^\ell & \textit{when} \quad a_\varphi^k \geqslant k+1 \textit{ or } a_\varphi^k = n; & \end{array}$$

(c) *judgment cancellation if and only if for all $\varphi \in \Phi$ and $k, \ell \in \{0, \ldots, n-1\}$,*

$$\begin{array}{llll} (1) & a_\varphi^{k+2} = a_\varphi^k + 2 & \textit{when} \quad a_\varphi^k \leqslant k \leqslant n-2 & \textit{and} \\ (2) & k+2 < a_\varphi^{k+2} & \textit{when} \quad a_\varphi^k \geqslant k+1 \textit{ or } a_\varphi^k \geqslant n-1. & \end{array}$$

Proof. The conditions follow from a careful analysis of the axioms.

[9] It is not generally true that independence implies monotonicity—yet, it is true for quota rules.

(a) The variable absolute quota rule F_{va} with thresholds a^k_φ is judgment monotonic if and only if, whenever we add extra support to a proposition φ in a profile with k reported judgments on $\widetilde{\varphi}$, (i) the new threshold on φ is reached in case the old threshold on φ was reached and (ii) the new threshold on $\neg\varphi$ is not reached if the old threshold on $\neg\varphi$ was not reached. To "add extra support" includes two cases here: First, the case where an agent accepts φ instead of $\neg\varphi$, and second, the case where an agent accepts φ instead of abstaining on $\widetilde{\varphi}$.

For the first case, where the number of reported judgments on $\widetilde{\varphi}$ remains the same, conditions (i) and (ii) are trivially satisfied.

For the second case, it is easy to see the following: If $a^{k+1}_\varphi \leqslant a^k_\varphi + 1$ whenever $a^k_\varphi \leqslant k \leqslant n-1$, then condition (i) will be satisfied, and if $a^k_\varphi \leqslant a^{k+1}_\varphi$ whenever $a^{k+1}_\varphi \leqslant k$ (specifically for proposition $\neg\varphi$), then condition (ii) will be satisfied.

On the other hand, if there exists $\psi \in \Phi$ such that $a^{k+1}_\psi > a^k_\psi + 1$ for $a^k_\psi \leqslant k \leqslant n-1$, we can construct a profile $\boldsymbol{J}$ where k agents report a judgment on $\widetilde{\psi}$ and a^k_ψ of them accept ψ (meaning that ψ will be collectively accepted on $\boldsymbol{J}$), and a different profile $\boldsymbol{J}'$ where $k+1$ agents report a judgment on $\widetilde{\psi}$ and $a^k_\psi + 1$ of them accept ψ (meaning that ψ will be collectively rejected on $\boldsymbol{J}'$), violating condition (i). Similarly, if there exists $\psi \in \Phi$ such that $a^k_\psi > a^{k+1}_\psi$ for $a^{k+1}_\psi \leqslant k$, we can construct a profile $\boldsymbol{J}$ where k agents report a judgment on $\widetilde{\psi}$ and a^{k+1}_ψ of them accept ψ (meaning that ψ will be collectively rejected on $\boldsymbol{J}$), and a different profile $\boldsymbol{J}'$ where $k+1$ agents report a judgment on $\widetilde{\psi}$ and still a^{k+1}_ψ of them accept ψ (that is, proposition $\neg\psi$ obtained extra support), meaning that ψ will be collectively accepted on $\boldsymbol{J}'$ and violating condition (ii).

(b) For the "if" direction: Consider two arbitrary profiles $\boldsymbol{J}$ and $\boldsymbol{J}'$ such that $n^{\boldsymbol{J}}_\varphi = n^{\boldsymbol{J}'}_\varphi$ and suppose that they have a different number of agents reporting a judgment on $\widetilde{\varphi}$ (otherwise independence holds trivially for the variable absolute quota rule $F_{a^k_\varphi}$). Take $n^{\boldsymbol{J}}_{\widetilde{\varphi}} = k$ and $n^{\boldsymbol{J}'}_{\widetilde{\varphi}} = \ell$, with $k < \ell$ (the case where $\ell > k$ is symmetric). We will show that the axiom of independence is satisfied with respect to $\boldsymbol{J}$ and $\boldsymbol{J}'$.

If $a^k_\varphi \leqslant k \leqslant n-1$, then $a^\ell_\varphi = a^k_\varphi$ (by condition (1)). This means that φ will be treated the same in the profiles $\boldsymbol{J}$ and $\boldsymbol{J}'$ and independence will be satisfied.

If $a^k_\varphi \geqslant k+1$ or $a^k_\varphi = n$, then by condition (2) we have that $k < a^\ell_\varphi$. In this case, whenever φ is collectively accepted in $\boldsymbol{J}$, independence is vacuoulsy satisfied (because $a^k_\varphi \geqslant k+1$ directly implies the rejection of φ, while when $a^k_\varphi = n$ and φ is accepted, there is no extra support to be added on $\neg\varphi$ in order to obtain the profile $\boldsymbol{J}'$). Now, whenever φ is collectively rejected in $\boldsymbol{J}$, it will hold that $k < a^k_\varphi$, so $a^k_\varphi \geqslant k+1$. Then, $n^{\boldsymbol{J}'}_\varphi = n^{\boldsymbol{J}}_\varphi \leqslant k < a^\ell_\varphi$, therefore φ will be rejected in the profile $\boldsymbol{J}'$ as well, and thus independence is satisfied.

For the "only if" direction: We will work on the contrapositive. We will show that if condition (1) or condition (2) does not hold, then we can have two profiles $\boldsymbol{J}$ and $\boldsymbol{J}'$ on which the axiom of independence is violated.

If condition (1) does not hold, then there are $\varphi \in \Phi$ and $k, \ell \in \{0, \dots, n\}$, $k < \ell$, with $a^k_\varphi \leqslant k \leqslant n-1$ such that $a^\ell_\varphi \neq a^k_\varphi$. If $a^\ell_\varphi < a^k_\varphi$, consider the following profiles: $\boldsymbol{J}$, where a^ℓ_φ agents accept φ and $k - a^\ell_\varphi$ agents accept $\neg\varphi$ (thus $n^{\boldsymbol{J}}_{\widetilde{\varphi}} = k$) and $\boldsymbol{J}'$, where a^ℓ_φ agents (the same as in $\boldsymbol{J}$) accept φ and $\ell - a^\ell_\varphi$ agents accept $\neg\varphi$ (thus $n^{\boldsymbol{J}'}_{\widetilde{\varphi}} = \ell$). Then, φ will be rejected in $\boldsymbol{J}$ but accepted in $\boldsymbol{J}'$, violating independence. If $a^k_\varphi < a^\ell_\varphi$, the construction is symmetric.

If condition (2) does not hold, then there are $\varphi \in \Phi$ and $k, \ell \in \{0, \dots, n\}$, $k < \ell$, with $a^k_\varphi \geqslant k+1$ or $a^k_\varphi = n$ such that $a^\ell_\varphi \leqslant k$. We construct a profile $\boldsymbol{J}$ with exactly k agents judging the issue $\widetilde{\varphi}$, and all of them accepting φ. Since $k < a^k_\varphi$, the proposition φ will be rejected in $\boldsymbol{J}$. Then, we construct another profile $\boldsymbol{J}'$ with exactly ℓ agents judging the issue $\widetilde{\varphi}$, and exactly k of them (the same as in $\boldsymbol{J}$) accepting φ. Since $k \geqslant a^\ell_\varphi$, proposition φ will be accepted in $\boldsymbol{J}'$. Thus, independence is violated.

(c) The proof is analogous to that of part (b). □

Figure 4.5 graphically presents the observations of this section.

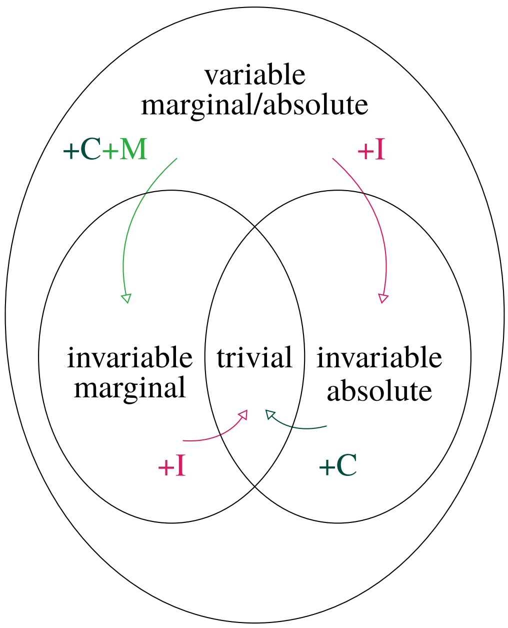


Figure 4.5: Quota rules for incomplete judgments and their properties.

4.3 On Positional Scoring Rules (Voting)

In the voting framework, we have thus far axiomatically examined only specific rules in the class of weight rules, in Section 4.1.2. This section solves an analogous exercise, but for the class of all positional scoring rules that we defined in Chapter 3, Section 3.3.

Smith (1973) and Young (1975) initiated the axiomatic analysis of scoring rules given complete preferences. Two decades later, Myerson (1995) generalised the previous results to profiles of votes that could take any form over the set of alternatives, and it is this latter work that we draw upon here.

> The characterisation results that we provide require an infinite superpopulation. This is a crucial assumption that enables us to construct profiles of arbitrary sizes, and, informally speaking, stir the outcome towards specific winning alternatives under given scoring functions. For the remainder of this section, we will thus omit referring to the size of the superpopulation, but the reader should keep in mind that a countably infinite one is presupposed. On the other hand, no assumption about the set of all potential alternatives is needed.

4.3.1 Characterising All Positional Scoring Rules

The following are obvious translations of axioms from the complete (Smith, 1973; Young, 1975) to the incomplete voting framework. The counterparts of anonymity and neutrality in judgment aggregation have also been discussed in Section 4.2.1.

> Our definitions and results are demonstrated for the most general case of preferences $R \in \mathcal{D}(A)$, which are acyclic and not necessarily transitive (also, not necessarily strict). But one could also consider strict or transitive preferences, as well as top-truncated preferences, without significant technical consequences.

Axiom: ***Anonymity.*** *For all permutations $\pi \colon N \to N$ and profiles $\boldsymbol{R} = (R_1, \dots, R_n)$, it holds that $F(\pi(\boldsymbol{R})) = F(\boldsymbol{R})$, where $\pi(\boldsymbol{R}) = (R_{\pi(1)}, \dots, R_{\pi(n)})$.*

Axiom: ***Neutrality.*** *For all permutations $\sigma \colon A \to A$ and profiles $\boldsymbol{R} = (R_1, \dots, R_n)$, it holds that $F(\sigma(\boldsymbol{R})) = \sigma(F(\boldsymbol{R}))$, where $\sigma(F(\boldsymbol{R})) = \{\sigma(a) \mid a \in F(\boldsymbol{R})\}$ and $\sigma(\boldsymbol{R}) = (\sigma(R_1), \dots, \sigma(R_n))$.*

Axiom: ***Reinforcement.*** *For all profiles $\boldsymbol{R}$ and $\boldsymbol{R}'$, if $F(\boldsymbol{R}) \cap F(\boldsymbol{R}') \neq \emptyset$, then $F(\boldsymbol{R}, \boldsymbol{R}') = F(\boldsymbol{R}) \cap F(\boldsymbol{R}')$.*

Axiom: ***Continuity.*** *For all profiles $\boldsymbol{R}$ and $\boldsymbol{R}'$, there exists a positive integer K such that, for every integer k that is greater than K, it holds that $F(\underbrace{\boldsymbol{R}, \dots, \boldsymbol{R}}_{k}, \boldsymbol{R}') \subseteq F(\boldsymbol{R})$.*

Anonymity requires that the outcome of the aggregation should not depend on the names of the agents but only on the preferences they contribute; neutrality says that all alternatives should be treated symmetrically; reinforcement prescribes that if two groups unite and vote together, then the alternatives that win should be the alternatives that would win for both elections if each group were to vote separately (unless there are no such alternatives); finally, continuity states that a sufficiently large number of agents should be able to change the outcome in accordance with their preference.

Of these axioms, anonymity, reinforcement, and continuity hold for scoring rules, and only those (Myerson, 1995). This leaves neutrality, which is to be connected to positional scoring functions. However, non-positional and positional scoring functions may be equivalent.[10] Because of this, non-positional scoring functions may define neutral scoring rules. Let us define a positional scoring *rule* as one that can be represented by *some* positional scoring function. Lemma 4.32 helps bridge the gap between positional scoring rules and positional scoring functions, and leads into the axiomatisation (Theorem 4.33).

4.32. Lemma. *A scoring function s is equivalent to some positional scoring function if for all alternatives $x, y \in A$ and preferences $R \in \mathcal{D}(A)$, two conditions hold:*

1. $s_R(x) - s_R(y) = s_{R_{\langle xy \rangle}}(y) - s_{R_{\langle xy \rangle}}(x)$
2. $s_R(z) - s_R(y) = s_{R_{\langle xy \rangle}}(z) - s_{R_{\langle xy \rangle}}(x)$*, for all $z \neq x, y$*

Proof. By the fact that all permutations over A can be expressed as successive permutations only involving two alternatives, (1) and (2) together imply that, for an arbitrary permutation function σ, preference R, and alternatives a and b,

$$s_R(a) - s_R(b) = s_{R_\sigma}(a_\sigma) - s_{R_\sigma}(b_\sigma) \tag{4.1}$$

Consider $|A|$ applications of Equation (4.1) to an arbitrary $a \in A$:

$$\sum_{x \in A} s_R(a) - s_R(x) = \sum_{x \in A} s_{R_\sigma}(a_\sigma) - s_{R_\sigma}(x_\sigma)$$

This can be rewritten as follows:

$$|A| \cdot s_R(a) - \sum_{x \in A} s_R(x) = |A| \cdot s_{R_\sigma}(a_\sigma) - \sum_{x \in A} s_{R_\sigma}(x_\sigma)$$

So, given a scoring function s' that satisfies (1) and (2), in order to show that $s'_R(a) = s'_{R_\sigma}(a_\sigma)$ it suffices to show the following:

$$\sum_{x \in A} s'_R(x) = \sum_{x \in A} s'_{R_\sigma}(x_\sigma) \tag{4.2}$$

Suppose that (1) and (2) hold for s. We want to find an equivalent positional scoring function s'. If s is positional, we are done because it is equivalent to itself. If not, there are some R^*, some permutation function σ, and some alternative x such

[10]For example, take $a, b \in A$: For aRb let $s_R(a) = 1$, $s_R(b) = 0$; for $bR'a$, let $s_{R'}(a) = 2$, $s_{R'}(b) = 1$; set all other scores to 0.

that $s_{R^*}(x) \neq s_{R^*_\sigma}(x_\sigma)$. Fix an alternative $a \in A$. Define s' as follows: First set $s'_{R^*}(x) = s_{R^*}(x)$. Next, for each distinct permutation R^*_ρ of R^*, define the following:

$$s'_{R^*_\rho}(y) = s_{R^*_\rho}(y) + s_{R^*}(a) - s_{R^*_\rho}(a_\rho) \quad \text{for all } y \in A \tag{4.3}$$

After going through the permutations of R^*, there may still be some other $R^\dagger$, ω and y such that $s_{R^\dagger}(y) \neq s_{R^\dagger_\omega}(y_\omega)$, in which case the process needs to be repeated for $R^\dagger$. As there are only finitely many preferences and the process will not consider the same preference twice, eventually there will be no more such cases. For all remaining preferences R set $s'_R = s_R$. Note that s' is equivalent to s and that (1) and (2) still hold.

We have above defined s'. Now take arbitrary $x \in A$, preference R, and permutation function σ. We want to show that $s'_R(x) = s'_{R_\sigma}(x_\sigma)$. If the score s was changed for no permutation of R, then $s_R(x) = s_{R_\sigma}(x_\sigma)$ and indeed $s'_R(x) = s'_{R_\sigma}(x_\sigma)$ as required. Otherwise, there are some R^* for which $s_{R^*} = s'_{R^*}$, and permutation functions ρ, ω such that $R^*_\rho = R$ and $R^*_\omega = R_\sigma$. We can suppose that these are those permutation functions that were used in the construction of s'—if either is instead the identity permutation, the required equality trivially holds. Regardless, Equation (4.2) holds:

$$\begin{aligned}
\sum_{x\in A} s'_R(x) &= \sum_{x\in A} s'_{R^*_\rho}(x_\rho) && \text{because alternatives appear exactly once}\\
&= \sum_{x\in A} s_{R^*_\rho}(x_\rho) + s_{R^*}(a) - s_{R^*_\rho}(a_\rho) && \text{by Equation (4.3)}\\
&= \sum_{x\in A} s_{R^*}(x) && \text{by Equation (4.1)}\\
&= \sum_{x\in A} s'_{R^*}(x) && \text{by definition of } s'
\end{aligned}$$

The equality concerning R^* and $R^*_\omega = R_\sigma$ is obtained in an identical manner. □

Before proceeding to the main characterisation of positional scoring rules by means of anonymity, neutrality, reinforcement, and continuity, it is worthwhile to wonder whether there are other nicely defined rules for incomplete preferences that violate some of these axioms.[11] For instance, consider a majority-based rule that selects as winners those alternatives that are deemed superior to the largest number of other alternatives by a strict majority of agents, which may be regarded as a generalisation of the *Copeland* rule (Zwicker, 2016). Suppose we have the following profile with three agents:

[11]Recall that the weight rules of winner type that we have seen in Section 3.1.2 belong to the class of positional scoring rules, and will thus satisfy our axioms.

$$
\begin{tikzcd}[column sep=small]
 & a \arrow[d] & \\
 & b \arrow[d] & \\
 & d \arrow[dl] \arrow[dr] & \\
c & & e
\end{tikzcd}
\qquad
\begin{tikzcd}[column sep=small]
 & c \arrow[d] & \\
 & a \arrow[d] & \\
 & b \arrow[dl] \arrow[dr] & \\
d & & e
\end{tikzcd}
\qquad
\begin{tikzcd}[column sep=small]
 & b \arrow[d] & \\
 & c \arrow[d] & \\
 & a \arrow[dl] \arrow[dr] & \\
d & & e
\end{tikzcd}
$$

In the above profile, the winning set is $\{a, b, c\}$, since for each of these alternatives there exist two agents that rank it higher than exactly three other alternatives.

The reinforcement axiom is violated by this rule: Add to the given profile the single-agent profile consisting of the complete preference $a \rhd b \rhd c \rhd d \rhd e$, where our rule would output alternative a as the unique winner. Although reinforcement demands that a is the only winner in the combined profile too, a and b are both considered superior to three other alternatives by a strict majority of agents and thus must both win.

4.33. Theorem. *A voting rule for incomplete preferences is a positional scoring rule if and only if it satisfies anonymity, neutrality, reinforcement, and continuity.*

Proof. Every positional scoring rule obviously satisfies all the axioms of the statement, so the "only if" holds. For the "if": a voting rule over incomplete preferences that satisfies anonymity, neutrality, reinforcement, and continuity has to be a scoring rule F_s for some scoring function s by Myerson (1995). If s is trivial, then F_s is positional. For non-trivial s, we show that (1) and (2) of Lemma 4.32 hold, and thus that the scoring rule F_s is positional.

For (1), fix arbitrary alternatives $a, b \in A$ and preference R. We will construct a profile where both a and b are winning, and such that if (1) did not hold they could not have the same score. For our construction we require a profile where a and b have the same summed score which is arbitrarily larger than the score of any other alternative. To do this, we first need to show the following:

$$\text{there exists a profile } \boldsymbol{R} \text{ such that } 2 \leqslant |F_s(\boldsymbol{R})| < |A| \tag{4.4}$$

Since s is not trivial, there is some R' such that for nonempty $X \subsetneq A$, it holds that $s_{R'}(x) = s_{R'}(y) > s_{R'}(z)$ for all $x, y \in X$, $z \in A \backslash X$. If $|X| > 1$ for some such set X, we have the required profile of (4.4). So suppose that $X = \{x\}$ for all relevant X. Pick y such that $s_{R'}(y) \geqslant s_{R'}(z)$ for all $z \neq x$. It is true by neutrality that y gets the unique highest score on $R'_{\langle xy \rangle}$ and that in the profile $(R', R'_{\langle xy \rangle})$, x and y must be joint winners with the same summed score. So in this case we also have the required profile of (4.4).

Now, (4.4) provides a profile $\boldsymbol{R}$ where some distinct x and y are winning and some z is not winning. Let Ω be the set of all permutations where x and y are fixed points. In the profile $(\omega(\boldsymbol{R}), \omega(\boldsymbol{R}_{\langle xy \rangle}))_{\omega \in \Omega}$ the alternatives x and y are the unique winners. To see this, take some alternative $z \neq x, y$, and note that, by neutrality, for each profile $\omega(\boldsymbol{R})$, alternative x gets score at least as high as the score of z, and for some such profile x gets strictly higher score than z; the same applies for $\omega(\boldsymbol{R}_{\langle xy\rangle})$, so z does not win; finally by neutrality x cannot win without y winning. Permute a with x and b with y. Enough copies of this profile create an arbitrarily large gap between the score of a and b and the scores of other alternatives (and permuting a and b does not change this new profile).

Let t be the maximal difference between the score of two alternatives in R and between the score of two alternatives in $R_{\langle ab\rangle}$. Formally, we define t as follows:

$$t = \max\left(\max_{x,y\in A} s_R(x) - s_R(y), \max_{x,y\in A} s_{R_{\langle ab\rangle}}(x) - s_{R_{\langle ab\rangle}}(y)\right)$$

There is a profile $\boldsymbol{R}'$ where a and b get (the same) summed score at least $2t+1$ greater than all the other alternatives and $\boldsymbol{R}'_{\langle ab\rangle} = \boldsymbol{R}'$. Let us add R and $R_{\langle ab\rangle}$ to $\boldsymbol{R}'$. By construction, no alternative other than a and b can have the maximal summed score. By neutrality, both a and b must be winning. But if (1) did not hold, only one of them would be winning. Thus part (1) of Lemma 4.32 holds.

The general procedure for part (2) of Lemma 4.32 is similar; though we must also start with arbitrary $c \in A \setminus \{a, b\}$. As before, there is a profile $\boldsymbol{R}$ such that $F_s(\boldsymbol{R}) = \{a, b\}$ and $\boldsymbol{R}_{\langle ab\rangle} = \boldsymbol{R}$. We define $\boldsymbol{R}' = (\boldsymbol{R}_{\langle ac\rangle}, \boldsymbol{R}_{\langle bc\rangle})$, satisfying the following:

$$F_s(\boldsymbol{R}') = \{c\} \qquad \text{and} \qquad \boldsymbol{R}'_{\langle ab\rangle} = \boldsymbol{R}'$$

Note further that a and b receive the same score under $\boldsymbol{R}'$. Next define $\boldsymbol{R}'' = (\boldsymbol{R}, \boldsymbol{R}_{\langle ac\rangle}, \boldsymbol{R}_{\langle bc\rangle})$, for which the following holds:

$$F_s(\boldsymbol{R}'') = \{a, b, c\} \qquad \text{and} \qquad \boldsymbol{R}''_{\langle ab\rangle} = \boldsymbol{R}''$$

Suppose that (2) does not hold. Suppose in particular that the following is the case:[12]

$$s_R(c) - s_R(b) > s_{R_{\langle ab\rangle}}(c) - s_{R_{\langle ab\rangle}}(a)$$

Roughly speaking, c performs "better" against b under R than c performs against a under $R_{\langle ab\rangle}$. For an integer k, write $k.R$ for k copies of R and $k.\boldsymbol{R}$ for k copies of $\boldsymbol{R}$. We can choose suitable integers j, k, ℓ, and m such that we ensure the following:

- $F_s(j.R, k.\boldsymbol{R}, l.\boldsymbol{R}', m.\boldsymbol{R}'') = c$
- the score of a is greater than the score of c in $(j.R_{\langle ab\rangle}, k.\boldsymbol{R}, l.\boldsymbol{R}', m.\boldsymbol{R}'')$

But we have that $(j.R, k.\boldsymbol{R}, l.\boldsymbol{R}', m.\boldsymbol{R}'')_{\langle ab\rangle} = (j.R_{\langle ab\rangle}, k.\boldsymbol{R}, l.\boldsymbol{R}', m.\boldsymbol{R}'')$, which stands in contradiction with neutrality. □

[12]The case of the inverse strict inequality is symmetric.

4.3.2 Characterising Generalisations of the Borda Rule

Having singled out positional scoring rules amongst all aggregation rules for incomplete preferences, we are prepared to move on to a subclass of special interest, viz. the Borda class for top-truncated preferences (recall Section 3.3.2).

Our analysis will proceed in steps, each of which will concern a smaller class of rules than the previous one: First, we are going to add some new axioms to the characterisation of all positional scoring rules for top-truncated preferences, and obtain all positional scoring rules in the Borda class. Then, by considering a few further axioms, we will be able to restrict attention to our specific generalisations of the Borda rule within the Borda class.

Our axioms will provide a principled way of understanding not only the differences, but also the similarities of the suggested generalisations of the Borda rule. The axiomatic characterisation of the Borda class is based on the original characterisation of the Borda rule for domains of linear orders by Young (1974). In his characterisation, Young used the following four axioms: *neutrality*, *reinforcement*, *faithfulness*, and *cancellation* (which we will formally define later on), with the last being the most critical one for the identification of the Borda rule. Analogously, essential for the characterisation of the Borda class will be our axiom of *top-cancellation*. Top-cancellation extends Young's cancellation axiom to top-truncated orders, by requiring that the voting rule should not distinguish between alternatives that are strictly ranked by all agents and tie in pairwise comparisons.

On the other hand, each specific rule within the Borda class satisfies distinctive properties, which are brought to the surface when expressed formally as axioms. For instance, we will show that by enforcing Young's original cancellation axiom within the Borda class, we can specify the averaged Borda rule. This result is also in line with the work of Cullinan et al. (2014), who have characterised this specific version of the Borda rule for domains of partial orders, relying on the four classical axioms of Young. The optimistic and the pessimistic Borda rules will be obtained by imposing two axioms that are reminiscent of monotonicity.

Although our work is tightly connected to the characterisation of the Borda rule by Young (1974), other characterisations of the Borda rule in the same formal setting have also been produced by Farkas and Nitzan (1979) and by Saari (1990). The former have used the axiom of Pareto stability based on a notion of relative unanimity, while the latter has employed weaker versions of Young's axioms and has incorporated the axiom of anonymity as well.

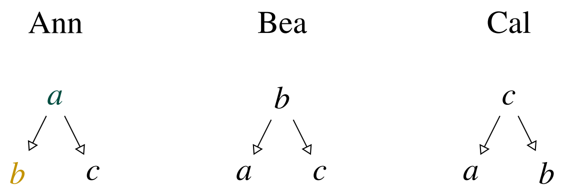
We also note that aggregation processes based on the Borda scores, together with their corresponding axiomatic properties, have received much attention in several settings beyond voting as well. Nitzan and Rubinstein (1981) have characterised the Borda rule as a social welfare function (i.e., a function that outputs social rankings instead of winning alternatives). Duddy et al. (2016) and Brandl and Peters (2019) have focused on aggregation mechanisms that produce collective dichotomous preferences and are inspired by Borda's form of scoring.

Characterising the Borda Class

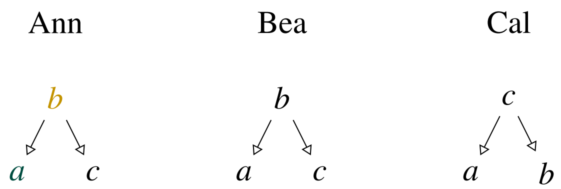
We will provide two distinct characterisations of the Borda class of rules for top-truncated preferences. Recall that the Borda class only includes positional scoring rules for which it is the case that, if a is strongly preferred to b (i.e., $a{>}b$), then the score of a is larger than the score of b in $\succsim$ (i.e., $s_{\succsim}(a) > s_{\succsim}(b)$). This is formally imposed by a *monotonicity* property, intuitively prescribing that moving an alternative to a "higher" position in a preference is beneficial to that alternative.

Axiom: ***Monotonicity.*** *Consider any preference $\succsim_i$ within any profile $\boldsymbol{\succsim}$ and two alternatives a, b such that $a{>}_i b$. If $b \in F(\boldsymbol{\succsim})$, then $F(\boldsymbol{\succsim}_{-i}, \succsim_i^{\langle ab \rangle}) = \{b\}$.*

4.34. Example. Take a profile of three agents as follows, and suppose that the winning alternatives under some rule F are a, b and c:



Then, consider a new profile where Ann swaps the position of a and b:



Monotonicity dictates that now b should be the only winner. $\triangle$

> In Section 4.2.1, we defined two monotonicity axioms for judgment aggregation. The essence of those axioms is akin to the monotonicity property we just introduced for voting. However, an important difference exists: In voting, it makes sense to demand that an increase on an alternative's support breaks a tie in favour of that alternative over all other alternatives (as implied by the axiom above); in judgment aggregation, propositions are not in competition with each other, so such a tie-breaking requirement is unreasonable.

Lemma 4.35 makes the link between monotonicity of a positional scoring rule F_s and the scoring function s defining that rule precise. Note also that this statement (here only described for the two top-truncated domains) holds for wider preference domains as well, like any kind of preorders.

4.35. Lemma. *A positional scoring rule F_s for top-truncated preferences satisfies monotonicity if and only if $a{>}b$ implies that $s_{\succsim}(a) > s_{\succsim}(b)$.*

Proof. To see that monotonicity is satisfied by any positional scoring rule for which the required condition on the scores holds, note that by flipping the positions of a and b in $\succsim$, the score of b increases, the score of a decreases, and all other scores remain the same (by definition of the positional scoring rule). Thus, if b was among the winners before the flipping, then afterwards it will be the unique winner with the highest score.

For the other direction, given a preference $\succsim$ and alternatives a, b with $a{>}b$, we can construct a profile $\boldsymbol{\succsim}$ where a and b are amongst the winners under the rule F_s, with arbitrarily large score:

Consider a preference $\succsim^*$ and two positions in $\succsim^*$ such that the score of an alternative in the first position is at least as large as the score of an alternative in the second position, which is at least as large as the score of the alternatives in all other positions. Then, consider a preference $\succsim'$ where a takes the first position in $\succsim^*$ while b takes the second position in $\succsim^*$, and a preference $\succsim''$ where b takes the first position in $\succsim^*$ while a takes the second position in $\succsim^*$. With sufficiently many (and equally many) copies of these preferences, we ensure that we have a profile $\boldsymbol{\succsim}^*$ where a and b win with arbitrarily large score. Define the profile $\boldsymbol{\succsim} = (\boldsymbol{\succsim}^*, \succsim, \succsim^{\langle ab \rangle})$ and call i the agent that submits the preference $\succsim$ in $\boldsymbol{\succsim}$. Then, for alternative b to be the unique winner in the profile $(\boldsymbol{\succsim}_{-i}, \succsim^{\langle ab \rangle})$, the required condition on the scores must be satisfied. □

Our first characterisation of the Borda class is in line with the characterisation of the Borda rule by Young (1974), who—informally speaking—identified the Borda rule as the unique scoring rule that satisfies a *cancellation* property.[13] Here, we examine an axiom that is similar in flavour to Young's cancellation, but applies specifically to domains of top-truncated preferences. *Top-cancellation* concerns preference profiles $\boldsymbol{\succsim}$ with $TOP(\boldsymbol{\succsim}) \neq \emptyset$. In such profiles, if the agents' preferences between the top alternatives "cancel" each other, then no alternative can be considered better than the others in the top set, and hence all alternatives in that set should be treated the same by the voting rule. For domains of linear orders (that are special domains of top-truncated preferences), top-cancellation reduces to the standard cancellation axiom of Young.

Axiom: ***Top-cancellation.*** *Consider any profile $\boldsymbol{\succsim}$ with $TOP(\boldsymbol{\succsim}) \neq \emptyset$. Suppose that for all alternatives $x, y \in TOP(\boldsymbol{\succsim})$ with $x \neq y$, the following holds:*

$$|\{i \in N \mid x{>_i}y\}| = |\{i \in N \mid y{>_i}x\}|$$

Then, it must be the case that $TOP(\boldsymbol{\succsim}) \subseteq F(\boldsymbol{\succsim})$ or $TOP(\boldsymbol{\succsim}) \cap F(\boldsymbol{\succsim}) = \emptyset$.

[13] In voting with complete preferences, cancellation demands that whenever for any two alternatives a and b the same number of agents rank a above b and b above a, then all alternatives should be tied winners. The idea is that the power of a over b "cancels" the power of b over a. Recall that we have also defined an axiom named cancellation in the framework of judgment aggregation, but the two homonymous axioms are not formally related.

Note that top-cancellation allows for the case where no top alternative of a profile belongs to the winning set (i.e., where $TOP(\succsim) \cap F(\succsim) = \emptyset$). Indeed, top-cancellation is a weak axiom, only ensuring that all top alternatives will be treated symmetrically when appropriate conditions hold—that these alternatives should also be among the winners is a separate intuitive requirement, which we will later take care of with the axiom of monotonicity.

4.36. Theorem. *A voting rule for top-truncated preferences is in the Borda class if and only if it satisfies anonymity, neutrality, reinforcement, continuity, monotonicity, and top-cancellation.*

Proof. That all rules in the Borda class satisfy these properties is easy to see. For the other direction, we know that a voting rule F that satisfies anonymity, neutrality, reinforcement, and continuity must be a positional scoring rule by Theorem 4.33. Let us call this rule F_s. Moreover, F_s, satisfying neutrality, reinforcement, monotonicity, and top-cancellation, when restricted to profiles of preferences that are linear orders, must reduce to the Borda rule (Young, 1974).[14]

Now, consider an arbitrary top-truncated preference $\succsim$ (that is not a linear order), and a linear order L that extends $\succsim'$ (i.e., $\succsim' \subset L$), where $\succsim'$ is the same as $\succsim$ besides having its top alternatives placed in the reverse order. Figure 4.6 presents an illustration, where $TOP(\succsim) = \{a_1, \dots, a_{m-k}\}$ and $BOT(\succsim) = \{a_{m-k+1}, \dots, a_m\}$.

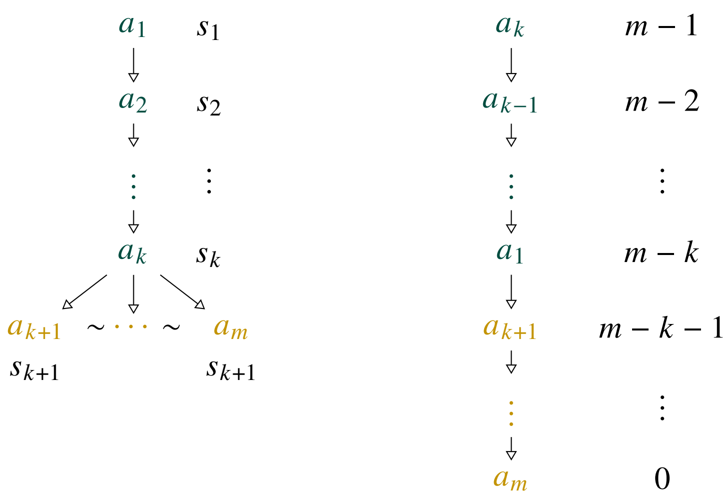


Figure 4.6: A top-truncated preference $\succsim$ (left) and a linearisation L of it with the order of its top alternatives reversed (right).

[14] Young's axiomatisation of the Borda rule relies on a *faithfulness* axiom, demanding that in any single-agent profile the winning alternative is the one on the top of that agent's preference. Straightforwardly, all monotonic positional scoring rules satisfy this property.

Next, construct the profile $\boldsymbol{\succsim} = (\succsim, L)$. Observe that $TOP(\boldsymbol{\succsim}) = TOP(\succsim) = TOP(L)$. From monotonicity and Lemma 4.35, we get $F_s(\boldsymbol{\succsim}) \subseteq TOP(\succsim)$, while top-cancellation entails that $TOP(\succsim) \subseteq F(\boldsymbol{\succsim})$ or $TOP(\succsim) \cap F(\boldsymbol{\succsim}) = \emptyset$. Thus, $F_s(\boldsymbol{\succsim}) = TOP(\succsim)$.

We know that the scores of the alternatives in L will be Borda-like. Moreover, the scores of all alternatives in $TOP(\boldsymbol{\succsim})$ have to be the same. So, we have the following:

$$\begin{aligned} & s_1 + m - k = s_2 + m - k + 1 = \ldots = s_{k-1} + m - 2 = s_k + m - 1 \\ \Rightarrow \quad & s_{\succsim}(a_1) = s_{\succsim}(a_2) + 1 = s_{\succsim}(a_3) + 2 = \ldots = s_{\succsim}(a_k) + k - 1 \end{aligned}$$

We conclude that F_s is in the Borda class. $\square$

Theorem 4.36—together with the characterisation of positional scoring rules presented in Theorem 4.33—implies the following corollary.

4.37. Corollary. *A positional scoring rule for top-truncated preferences is in the Borda class if and only if it satisfies monotonicity and top-cancellation.*

An immediate question that arises is whether the axioms appearing in Theorem 4.36 are all necessary for the characterisation result, i.e., whether they are independent. We know from the characterisation of positional scoring rules that anonymity, neutrality, reinforcement, and continuity are independent, and it is easy to see that monotonicity does not break this fact. Adding top-cancellation also preserves independence between these axioms. Proposition 4.38 states exactly this, and a proof for the most interesting case (i.e., that top-cancellation, together with the rest of our relevant axioms, does not imply anonymity) is provided. Note that this is a rather unexpected result, since—as we will see later too—in the original proof of Young (1974) the analogous cancellation axiom (in combination with the other axioms) ends up implying anonymity.

4.38. Proposition. *The axioms of anonymity, neutrality, reinforcement, continuity, monotonicity, and top-cancellation are logically independent on domains of top-truncated preferences.*

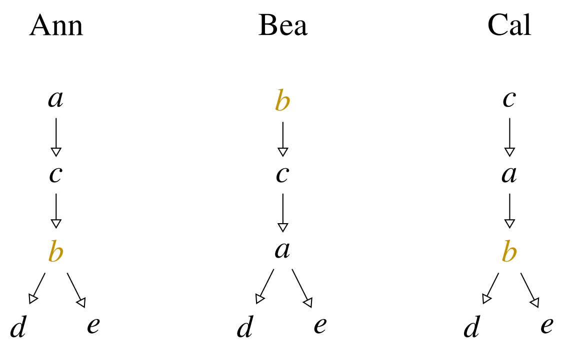
Proof. Let us show the most interesting case, i.e., that there exists a voting rule on top-truncated preferences that satisfies neutrality, reinforcement, continuity, monotonicity, and top-cancellation, but violates anonymity (analogous counterexamples can be easily found for all other combinations of our axioms as well).

Consider the voting rule F on top-truncated preferences that works just like the optimistic Borda rule, but with a small exception: when agent 1 reports a preference $\succsim^*$ that identifies some alternative $a \in A$ (whichever that a is) as the unique top one (that is, a preference $\succsim^* = \{(a, x) \mid x \in A \setminus \{a\}\}$), then agent 1 gets assigned double the weight of the other agents (i.e., the scores associated with her preference are twice the standard scores of the optimistic Borda rule). Obviously, this rule is not anonymous.

Neutrality is satisfied by F because the definition of the rule does not distinguish between names of alternatives. Monotonicity is also satisfied, because it is always better for an alternative to appear in a higher position in a preference. Continuity holds too, since by adding a sufficiently large number of copies of the same profile $\succsim'$ to a given profile $\succsim$, we can arbitrarily increase the Borda scores of the alternatives that win in $\succsim'$, and thus obtain the result prescribed by the axiom. To see that top-cancellation is satisfied, note that for any profile $\succsim$ with $TOP(\succsim) \neq \emptyset$, if the preference $\succsim^*$ defined above (for some given alternative a) appears in $\succsim$, it must be the case that $TOP(\succsim) = \{a\}$, and thus $F(\succsim) = \{a\}$ (which means that top-cancellation is vacuously satisfied). Otherwise, F functions as the standard optimistic Borda rule, and hence top-cancellation holds. Finally, F satisfies reinforcement: for all profiles $\succsim$ and $\succsim'$ and for any alternative $a \in A$, we have that the score that a receives by the rule F in the joint profile $(\succsim, \succsim')$ will always be the sum of the scores that a receives in $\succsim$ and in $\succsim'$. □

Our second characterisation of the Borda class relies on a result of Fishburn and Gehrlein (1976) for domains of linear orders (based on a proof sketch by Smith, 1973), namely that the Borda rule is the only positional scoring rule for which a Condorcet loser (CL) never wins (we call this property CL-consistency). We remind the reader that the Condorcet loser of a preference profile is an alternative that loses in pairwise comparisons to all other alternatives, where "losing" means that a majority of agents considers that alternative inferior to the one it is compared to. We extend this fundamental principle to profiles of top-truncated preferences by stipulating no alternative that is a Condorcet loser relative to the top part of a profile should ever be amongst the winners.[15]

4.39. Example. Consider a profile of three agents, where the top alternatives are a, b, and c, and Ann and Cal rank alternative b below every other alternative:



Top-CL-consistency implies that alternative b cannot be among the winners. △

Axiom: ***Top-CL-consistency.*** *Consider any profile $\succsim$ with $TOP(\succsim) \neq \emptyset$, and suppose that the following holds for some $b \in TOP(\succsim)$ and for all $x \in TOP(\succsim) \setminus \{b\}$:*

$$|\{i \in N \mid x \succ_i b\}| > |\{i \in N \mid b \succ_i x\}|$$

Then, it must be the case that $b \notin F(\succsim)$.

[15] Our axiom does not require that alternatives in the bottom part of a profile must be barred from winning as well (but this of course would be enforced by imposing monotonicity).

4.40. Theorem. *A voting rule for top-truncated preferences is in the Borda class if and only if it satisfies anonymity, neutrality, reinforcement, continuity, monotonicity, and top-CL-consistency.*

Proof. The proof is analogous to that of Smith (1973). □

Theorem 4.40, together with the characterisation of positional scoring rules, implies the following corollary.

4.41. Corollary. *A positional scoring rule for top-truncated preferences is in the Borda class if and only if it satisfies monotonicity and top-CL-consistency.*

At this point we also need to examine whether the axioms of Theorem 4.40 are independent. For example, can we find a positional scoring rule for top-truncated preferences that satisfies top-CL-consistency but is not monotonic? Proposition 4.42 answers this question in the affirmative.

4.42. Proposition. *The axioms of anonymity, neutrality, reinforcement, continuity, monotonicity, and top-CL-consistency are logically independent on domains of top-truncated preferences.*

Proof. We will show that there exists a positional scoring rule (a rule satisfying anonymity, neutrality, reinforcement, and continuity) for top-truncated preferences that satisfies top-CL-consistency but violates monotonicity. Showing independence for the remaining sets of axioms can be done in an analogous manner.

Consider the positional scoring rule for which, for an any given top-truncated preference $\succsim$, the Borda-like scores are assigned to the alternatives in $TOP(\succsim)$, and score $m+1$ is assigned to the alternatives in $BOT(\succsim)$. Clearly, this rule is not monotonic, but top-CL-consistency holds: no Condorcet loser within the top alternatives of a profile can ever win (either because an alternative that Pareto dominates it wins, or because some bottom alternative wins instead). □

Top-CL-consistency ensures that the scores of the top alternatives in an agent's top-truncated preference will be distributed in a linear manner, as required for rules in the Borda class. Note that classical CL-consistency—although applicable to domains of top-truncated preferences as well—is not appropriate for our purposes. In particular, not all rules in the Borda class satisfy CL-consistency. For example, consider the pessimistic Borda rule, and a profile with nine agents and four alternatives such that: four agents rank alternative a on top and every other alternative directly below, and the remaining five agents have preferences as follows: $b \succ_1 c \succ_1 d \succ_1 a$, $b \succ_2 c \succ_2 d \succ_2 a$, $c \succ_3 d \succ_3 b \succ_3 a$, $d \succ_4 b \succ_4 c \succ_4 a$, $d \succ_5 c \succ_5 b \succ_5 a$. Alternative a is the Condorcet loser of this profile, but it will be the winner according to the pessimistic Borda rule (it will receive 12 points, while all other alternatives will only get 10 points).

To sum up, the axioms of top-cancellation and of top-CL-consistency (together with monotonicity) are the ones that distinguish rules in the Borda class from all other positional scoring rules. Interestingly, these two axioms only bite for profiles of top-truncated preferences with a non-empty set of top alternatives. On the one hand, such profiles are rare in general. On the other hand, every preference can appear in some profile of that form. The key idea behind our proofs is that the rules with which we work are positional scoring rules. So, the score assigned to an alternative a in a given preference $\succsim$ will be fixed, and can be deduced by applying the relevant axioms in profiles with a non-empty set of top alternatives that $\succsim$ is part of.

Characterising Specific Rules in the Borda Class

After characterising the Borda class of rules via a number of normative axiomatic properties, in the remainder of this section we will identify those properties that charaterise each one of our specific rules of interest, within the Borda class.

We observe that the pessimistic Borda rule is the only rule in the Borda class for which the scores of the bottom alternatives in the top-truncated preferences do not depend on *how many* of these alternatives there are. Loosely speaking, this translates into the following slogan:

> The number of other alternatives with which some alternative a shares the bottom position does not affect a's performance.

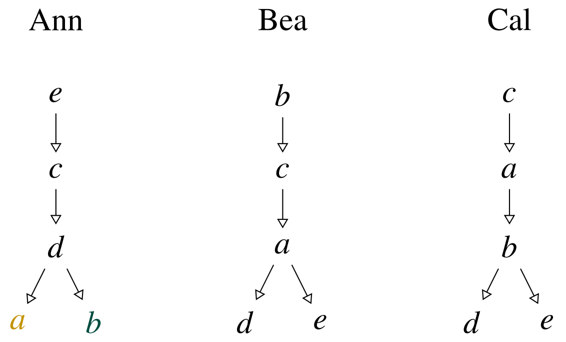
The axiom of *bot-indifference* formally captures this idea.

Axiom: ***Bot-indifference.*** *Consider two profiles $\boldsymbol{\succsim}$ and $\boldsymbol{\succsim}' = (\boldsymbol{\succsim}_{-i}, \succsim_i')$ for some agent i such that $\succsim_i'$ is obtained from $\succsim_i$ by having one of the bottom alternatives of $\succsim_i$, say alternative a, moved to the last position of the ranked alternatives in the top part. If $a \notin F(\boldsymbol{\succsim}')$, then for any $b \in BOT(\succsim_i) \setminus \{a\}$ the following holds:*

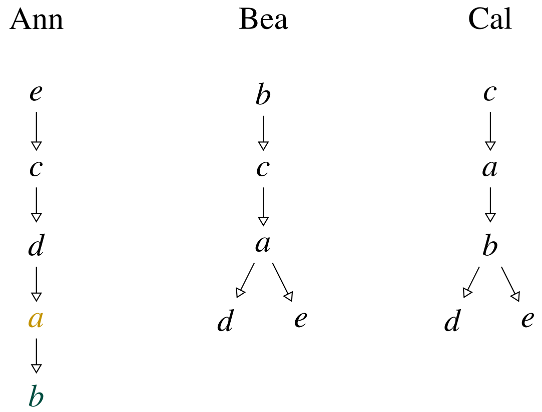
$$b \in F(\boldsymbol{\succsim}) \text{ if and only if } b \in F(\boldsymbol{\succsim}')$$

Thus, by moving a we create two profiles where the number of alternatives with which b shares a bottom position changes; we stipulate that this should not affect whether or not b is amongst the winners (at least not in case a is not winning in the second profile, the one where it intuitively is put in a better position).

4.43. Example. Consider a profile where Ann places alternatives a and b at the bottom of her preference, as follows:



Then, suppose that Ann moves alternative a to the lowermost position of her ranked alternatives above, leaving b to the bottom:



If a is not winning in the new profile, bot-indifference implies that alternative b is a winner in the former profile if and only if it is a winner in the latter one. △

As suggested earlier, bot-indifference characterises the pessimistic Borda rule.

4.44. Theorem. *The only voting rule for top-truncated preferences that simultaneously satisfies anonymity, neutrality, reinforcement, continuity, monotonicity, top-cancellation (or top-CL-consistency), and bot-indifference is the pessimistic Borda rule.*

Proof. We can easily verify that the pessimistic Borda rule satisfies all relevant axioms. For the other direction, suppose that we have a rule F_s, with corresponding scoring function s, in the Borda class (satisfying anonymity, neutrality, reinforcement, continuity, monotonicity, top-cancellation and top-CL-consistency) for which bot-indifference holds. Take two arbitrary preferences $\succsim$ and $\succsim'$ such that $\succsim'$ is obtained from $\succsim$ by having one of the bottom alternatives of $\succsim$, namely alternative a, moved to the last position of the ranked alternatives above. We will show that $s_{\succsim}(b) = s_{\succsim'}(b)$ for any alternative $b \in BOT(\succsim) \setminus \{a\}$. Note that in the special case where $\succsim'$ is a linear order, we know that b is the lowest alternative in $\succsim'$, and $s_{\succsim'}(b) = 0$. Thus, the bottom alternatives of all preferences $\succsim$ will be assigned score 0, meaning that F_s must be the pessimistic Borda rule.

Take the alternatives $b \in BOT(\succsim) \setminus \{a\}$ and $c \in TOP(\succsim)$ and construct a profile $\boldsymbol{\succsim}^*$ where b and c are the only winners. That is:

$$F(\boldsymbol{\succsim}^*) = \{b, c\}$$

The profile $\boldsymbol{\succsim}^*$ can be constructed by taking equally many copies of two preferences, one with b followed by c making up the top part and one with c followed by b making up the top part (and the remaining alternatives making up the bottom part in both cases). By using sufficiently many copies, we can take the difference in score between b and c and the next-best alternative to be arbitrarily large.

Then consider the following profiles, where $F(\boldsymbol{\succsim}) = \{b, c\}$ and $F(\boldsymbol{\succsim}') \subseteq \{b, c\}$:

$$\boldsymbol{\succsim} = (\boldsymbol{\succsim}^*, \succsim, \succsim^{\langle bc \rangle}) \quad \text{and} \quad \boldsymbol{\succsim}' = (\boldsymbol{\succsim}^*, \succsim', \succsim^{\langle bc \rangle})$$

So, $a \notin F(\boldsymbol{\succsim}')$, and $\boldsymbol{\succsim}' = (\boldsymbol{\succsim}_{-i}, \succsim_i')$, with $\succsim_i = \succsim$ and $\succsim_i' = \succsim'$. Hence, we have the following:

$$s_{\boldsymbol{\succsim}}(b) = s_{\boldsymbol{\succsim}}(c) = s_{\boldsymbol{\succsim}'}(c)$$

But $b \in F(\boldsymbol{\succsim})$ implies that $b \in F(\boldsymbol{\succsim}')$ by bot-indifference.

It must then be the case that $s_{\boldsymbol{\succsim}'}(b) \geqslant s_{\boldsymbol{\succsim}'}(c) = s_{\boldsymbol{\succsim}}(b)$. So, $s_{\boldsymbol{\succsim}_{-i}}(b) + s_{\succsim'}(b) \geqslant s_{\boldsymbol{\succsim}_{-i}}(b) + s_{\succsim}(b)$. We conclude that $s_{\succsim'}(b) \geqslant s_{\succsim}(b)$. Proving the inverse inequality in a symmetric manner, we have that $s_{\succsim'}(b) = s_{\succsim}(b)$. □

We obtain an immediate corollary:

4.45. Corollary. *The only voting rule for top-truncated preferences in the Borda class that satisfies bot-indifference is the pessimistic Borda rule.*

We next define a new axiomatic property, building on the basic idea that if the dominance relationships between different winning alternatives remain unaltered, then no tie between these alternatives can be broken. In words, *dom-power* suggests that a winning alternative a can *only* break a tie between itself and a different winning alternative b by having its support against b strictly increased.

Note that the optimistic Borda rule is the only rule in the Borda class for which, in any top-truncated preference, the score of the last ranked alternative on top remains the same if that alternative "moves" to the bottom instead. The axiom ensuring this is precisely *dom-power*.

Axiom: ***Dom-power.*** *Consider any two profiles $\boldsymbol{\succsim}$ and $\boldsymbol{\succsim}' = (\boldsymbol{\succsim}_{-i}, \succsim_i')$ such that the preference $\succsim_i'$ is obtained from the preference $\succsim_i$ by having one of the bottom alternatives of $\succsim_i$ moved to the last position of the ranked alternatives in the top part. Then, for any $a \in TOP(\succsim_i)$, the following holds:*

$$a \in F(\boldsymbol{\succsim}) \text{ if and only if } a \in F(\boldsymbol{\succsim}')$$

4.46. Example. In Example 4.43 that we presented previously, we have two profiles where Ann has moved alternative a to the last position of the ranked alternatives in the top part of her preference. In that example, any alternative in the top part of Ann's preference (like alternative e, c, or d) will be a winner in the former profile if and only if it is a winner in the latter one. △

4.47. Theorem. *The only voting rule for top-truncated preferences that simultaneously satisfies anonymity, neutrality, reinforcement, continuity, monotonicity, top-cancellation (or top-CL-consistency), and dom-power is the optimistic Borda rule.*

Proof. We can easily verify that the optimistic Borda rule satisfies all relevant axioms. For the other direction, suppose that we have a rule F_s in the Borda class, induced by a suitable scoring function s (and satisfying anonymity, neutrality, reinforcement, continuity, monotonicity, top-cancellation and top-CL-consistency) for which dom-power holds. Take two arbitrary preferences $\succsim$ and $\succsim'$ such that $\succsim'$ is obtained from $\succsim$ by having one of the bottom alternatives of $\succsim$, namely alternative b, moved to the last position amongst the ranked alternatives above. We will show that $s_{\succsim}(b) = s_{\succsim'}(b)$; thus, F must be the optimistic Borda rule.

Take an alternative $a \in TOP(\succsim)$ and a profile $\boldsymbol{\succsim}^*$, where the following holds:

$$F(\boldsymbol{\succsim}^*) = \{a, b\}$$

The profile $\boldsymbol{\succsim}^*$ can be easily constructed by taking equally many copies of two preferences, one with a followed by b making up the top part and one with b followed by a making up the top part (and the remaining alternatives making up the bottom part in both cases). By using sufficiently many copies, we can take the difference in score between a and b and the next-best alternative to be arbitrarily large.

Then consider the following profiles, where $F(\boldsymbol{\succsim}) = \{a, b\}$, $F(\boldsymbol{\succsim}') \subseteq \{a, b\}$, $\succsim_i = \succsim$, $\succsim_i' = \succsim'$, and $\boldsymbol{\succsim}' = (\boldsymbol{\succsim}_{-i}, \succsim_i')$:

$$\boldsymbol{\succsim} = (\boldsymbol{\succsim}^*, \succsim, \succsim^{\langle ab \rangle}) \quad \text{and} \quad \boldsymbol{\succsim}' = (\boldsymbol{\succsim}^*, \succsim', \succsim^{\langle ab \rangle})$$

So, the following holds:

$$s_{\boldsymbol{\succsim}}(b) = s_{\boldsymbol{\succsim}}(a) = s_{\boldsymbol{\succsim}'}(a)$$

Also, since $a \in F(\boldsymbol{\succsim})$, dom-power implies that $a \in F(\boldsymbol{\succsim}')$, meaning that the score of a in the profile $\boldsymbol{\succsim}'$ must be at least as high as the score of b. We hence have the following:

$$s_{\boldsymbol{\succsim}}(b) = s_{\boldsymbol{\succsim}'}(a) \geqslant s_{\boldsymbol{\succsim}'}(b)$$

But we know that $s_{\boldsymbol{\succsim}}(b) = s_{\boldsymbol{\succsim}_{-i}}(b) + s_{\succsim}(b)$ and $s_{\boldsymbol{\succsim}'}(b) = s_{\boldsymbol{\succsim}_{-i}}(b) + s_{\succsim'}(b)$, with which we can deduce that $s_{\succsim}(b) \geqslant s_{\succsim'}(b)$. The case with the inverse inequality can be proven symmetrically, and we conclude that $s_{\succsim}(b) = s_{\succsim'}(b)$. □

4.48. Corollary. *The only voting rule for top-truncated preferences in the Borda class that satisfies dom-power is the optimistic Borda rule.*

After realising that both the axioms of bot-indifference and of dom-power take the form of monotonicity-like properties, we easily see that they are independent of all other axioms in the characterisation of the Borda class.

We continue with the averaged Borda rule, which we are going to link to the property of *full-cancellation*. This axiom, in the spirit of top-cancellation, prescribes the equal status of all alternatives as far as the outcome of the aggregation process is concerned, and applies in cases where for all pairs of alternatives a, b the same number of agents prefers a to b and b to a.

Axiom: ***Full-cancellation.*** *Consider a profile* $\boldsymbol{\succsim} = (\succsim_1, \ldots, \succsim_n)$*, where the following holds for all* $x, y \in A$*:*

$$|\{i \in N \mid x \succ_i y\}| = |\{i \in N \mid y \succ_i x\}|$$

Then, it is the case that $F(\boldsymbol{\succsim}) = A$.

Note that full-cancellation reduces to the standard cancellation axiom for the special case of profiles of linear orders, and is in general logically independent of top-cancellation. Interestingly, when combined with other axioms that appear in the characterisation of the Borda class, full-cancellation becomes very strong:

4.49. Lemma. *Neutrality, reinforcement, monotonicity, and full-cancellation together imply anonymity, continuity, top-cancellation, and top-CL-consistency.*

Proof. First, it is easy to see that monotonicity (together with neutrality) implies that in every single-agent profile, that agent's top alternative is going to be the unique winner, which is the property of *faithfulness*. Then, in a similar manner to the one that Hansson and Sahlquist (1976) have used to prove that neutrality, reinforcement, faithfulness and (full-)cancellation characterise the Borda rule on domains of linear orders, we can show that these axioms directly characterise the rule represented by the symmetric Borda scores on top-truncated preferences, and therefore imply the rule in the Borda class satisfying anonymity, continuity, top-cancellation, and top-CL-consistency. □

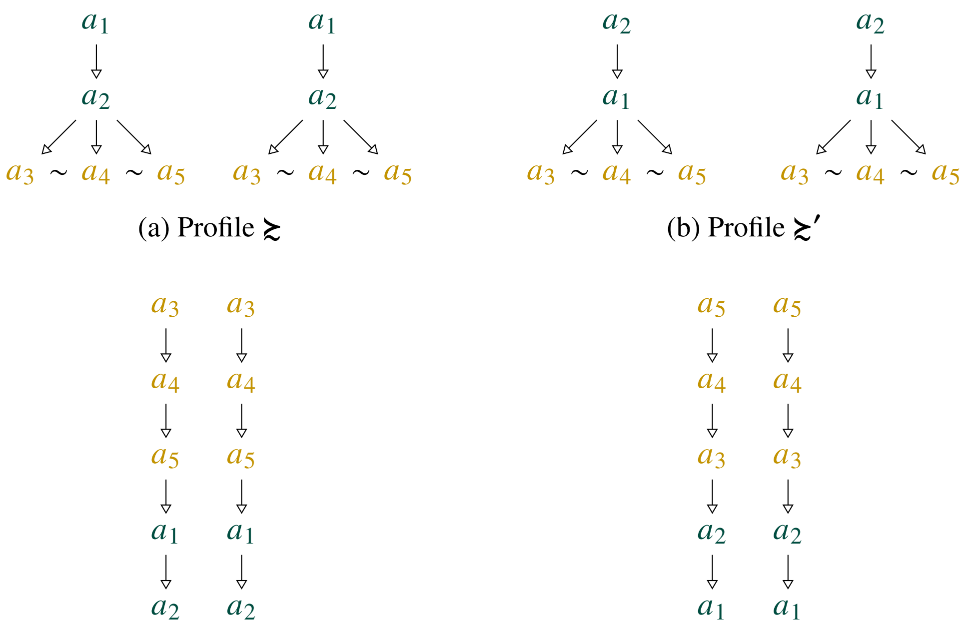
Using Lemma 4.49, we obtain a proof for the characterisation of the averaged Borda rule on top-truncated preferences (in Theorem 4.50) that explicitly hinges on the effect of full-cancellation within the Borda class. Although Theorem 4.50 could also be proven without any reference to the Borda class,[16] our proof sheds light on the particular way in which the averaged Borda rule differs from the optimistic and pessimistic Borda rules, by taking advantage of structurally analogous proof techniques.

[16]The fact that the characterisation of the rule represented by the symmetric Borda scores (i.e., the averaged Borda rule on top-truncated preferences) works for other domains of preorders, and for linear orders, is also informally mentioned by Young (1974) and by Hansson and Sahlquist (1976).

4.50. Theorem. *The only voting rule for top-truncated preferences that satisfies neutrality, reinforcement, monotonicity, and full-cancellation is the averaged Borda rule.*

Proof. After recalling that the averaged Borda rule corresponds to the symmetric way of defining domination scores (Proposition 3.34, in Section 3.3.2), it is not hard to see that this rule satisfies full-cancellation (and the other axioms of the statement).

For the other direction, take a rule that satisfies all the required axioms. By Lemma 4.49, we know that this rule also satisfies anonymity, continuity, top-cancellation, and top-CL-consistency, and hence is in the Borda class. Take such a rule F_s. Consider an arbitrary top-truncated preference $\succsim$ with $|TOP(\succsim)| = k$, with $1 \leqslant k \leqslant m-2$ (otherwise the proof is trivial). Let $\boldsymbol{\succsim}$ be the profile that consists of k copies of that preference $\succsim$, and let $\boldsymbol{\succsim}'$ be the profile that consists of k copies of the preference that reverses the order of the alternatives in the set $TOP(\succsim)$ and keeps the bottom alternatives in $\succsim$ unaltered. Then, we construct two profiles $\boldsymbol{L}$ and $\boldsymbol{L}'$ of linear orders such that $TOP(\boldsymbol{L}) = TOP(\boldsymbol{L}') = BOT(\succsim)$. Profile $\boldsymbol{L}$ consists of k preferences, all having the alternatives on top ranked in the same order (any arbitrary one); Profile $\boldsymbol{L}'$ also consists of k preferences, with the alternatives on top ranked in the same order, reversed from the one of $\boldsymbol{L}$. Moreover, in both $\boldsymbol{L}$ and $\boldsymbol{L}'$, every alternative on the bottom takes up each of the k positions exactly once. Figure 4.7 provides an example of this construction.



(a) Profile $\boldsymbol{\succsim}$ (b) Profile $\boldsymbol{\succsim}'$

(c) Profile $\boldsymbol{L}$ (d) Profile $\boldsymbol{L}'$

Figure 4.7: The construction of the profiles appearing in the proof of Theorem 4.50, for $m = 5$ and $k = 2$.

Then, consider the following profile $\boldsymbol{\succsim}''$, where full-cancellation applies:

$$\boldsymbol{\succsim}'' = (\boldsymbol{\succsim}, \boldsymbol{\succsim}', \boldsymbol{L}, \boldsymbol{L}')$$

Suppose that $TOP(\succsim) = \{a_1, \dots, a_k\} \subseteq A$ and $BOT(\succsim) = \{a_{k+1}, \dots, a_m\} \subseteq A$. By construction, we have the following conditions on the scores hold:

$$s_{\succsim''}(a_1) = \dots = s_{\succsim''}(a_k) = (m-1+m-k)k + \frac{k(k-1)}{2}2$$

$$s_{\succsim''}(a_{k+1}) = \dots = s_{\succsim''}(a_m) = 2ks_{k+1} + (m-1+k)k$$

But by full-cancellation we must have that $F_s(\succsim'') = A$, so all alternatives must have the same score. This means that the following holds:

$$(m-1+m-k) + (k-1) = 2s_k + (m-1+k) \quad \Leftrightarrow \quad s_{k+1} = \frac{m-k-1}{2}$$

We conclude that F_s is the averaged Borda rule. □

Theorem 4.50 and Lemma 4.49 imply the following corollary:

4.51. Corollary. *The only voting rule for top-truncated preferences in the Borda class that satisfies full-cancellation is the averaged Borda rule.*

Figure 4.8 wraps up this section, with a graphical summary of our results.

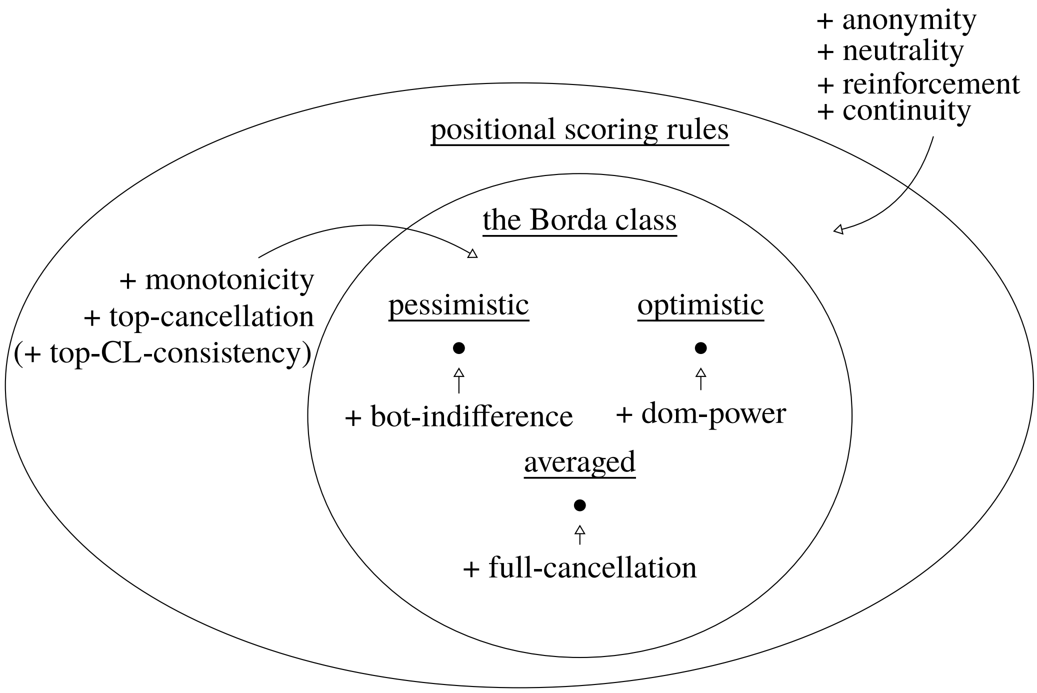


Figure 4.8: Positional scoring rules for incomplete preferences and their axioms (where the Borda class is defined for top-truncated preferences). The axioms next to an arrow indicate the properties required to obtain a class of rules, or a rule, within a larger class.

Chapter 5

Strategic Manipulation

The previous chapters of this thesis cover a wide range of topics on the aggregation of incomplete opinions, but all rely on a vital assumption: that every agent involved in a process of collective decision making reports sincere preferences or judgments. It is not hard to imagine that there are numerous situations where this is not the case. People driven by their personal interests make decisions together all the time, and artificial agents with individualised incentives may also engage in decision making within groups. Such intelligent agents are often expected to act in accordance with their desires rather than to be honest for the sake of virtue itself. Hence, new questions emerge, into which we will delve in this chapter:

> Are there settings where the best response of all agents is to be honest, no matter what opinions the other members of their group hold? Which aggregation rules fall victim to the strategic misrepresentation of the agents' individual opinions in the presence of incompleteness, and what kind of protection can we offer to prevent such behaviour?

Throughout this analysis, we will assume that all agents have full information about the aggregation rule in place and the opinions of their peers, as well as unlimited computational power to calculate the possible collective outcomes.

We set out to study strategic agents both in the framework of judgment aggregation (Section 5.1) and in the framework of voting (Section 5.2 and Section 5.3), keeping in mind contexts where incomplete opinions are central. To get a taste of the key challenges this chapter will bring out, let us revisit the weight rules in the framework of judgment aggregation that were introduced in Section 3.1.1.

5.1. Example. Consider the review phase of a conference involving three independent papers: p_1, p_2, and p_3. The sincere profile of judgments provided by our three reviewers is depicted in Table 5.1. The conference chair, being particularly convinced by the splitting axiom that characterises the equal-and-even weight rule (see page 29 for the definition of the rule and page 62 for the characterisation theorem), has decided to

	p_1	p_2	p_3
Ann:	☑	☑	☑
Bea:	☒	☑	☑
Cal:	☒	☑	☑
F_{ee}:	☒	☑	☑

Table 5.1: The sincere profile of judgments.

make use of that way of assigning weights to the submitted judgments of the reviewers. Since all sincere judgments are complete, the equal-and-even weight rule (as well as any other weight rule) will reduce to the standard majority rule, and will reject paper 1.

This outcome disagrees with Ann's judgment, who thinks that paper 1 should be accepted. Knowing the aggregation procedure that is implemented, Ann can manipulate the outcome by dishonestly abstaining on p_2 and p_3 (Table 5.2).

	p_1	p_2	p_3
Ann:	☑	☐	☐
Bea:	☒	☑	☑
Cal:	☒	☑	☑
F_{ee}:	☑	☑	☑

Table 5.2: The insincere profile of judgments.

Now, the support of Ann towards p_1 will get weight 1, while the stance of Bea and Cal against p_1 will weigh $\frac{1}{3} + \frac{1}{3}$ in total. Overall, Ann will manage to have paper 1 accepted without taking any risk regarding the other two papers. $\triangle$

Refraining from excessive formalities, we have seen that weight rules are very sensitive to the agents' strategic behaviour. In the remainder of this chapter, we will also inspect all other aggregation rules for incomplete preferences that have appeared in this thesis, with respect to their susceptibility to manipulation.

> In order to meticulously encode what it means for an agent to find a collective decision more preferable than a different one, we should clarify whether we are working with aggregation rules that are resolute (i.e., that always produce a single outcome) or not. If we are, as will be the case in Section 5.1 on judgment aggregation, then we simply have to define agents' preferences over single outcomes; if we are not, as will be the case within our voting framework, then we have to define preferences over sets of outcomes. The preferences of the agents regarding the collective decision should not be confused with their preferences over the specific alternatives under consideration.

5.1 Judgment Aggregation

The problem of *manipulation* was initially studied within judgment aggregation with complete inputs by Dietrich and List (2007c), and has been receiving increasing attention recently (Baumeister et al., 2013; Terzopoulou and Endriss, 2019c; Botan and Endriss, 2020). In this section, we offer a broader account, considering agents who have the freedom to report incomplete judgments, and may thus lie both by hiding their sincere judgment or by inventing a new dishonest judgment, as well as by flipping their judgment (e.g., from positive to negative) on an issue. Note that under the assumption of complete judgments, an agent can only lie by flipping her judgment on an issue.

In the context of single-issue voting with approval and participation quota, Maniquet and Morelli (2015) have been concerned with strategic agents that may insincerely abstain in order to obtain a more desirable outcome for themselves, and have remarked that those rules that satisfy the property of monotonicity prevent such behaviour. Although this observation will also be important for our work, Maniquet and Morelli have made further assumptions about the probabilistic information of the agents concerning the number of their peers that abstain, on which their results heavily rely. On the contrary, we follow a more basic—qualitative rather than quantitative—path and study cases where agents will *never* have the opportunity to benefit by being insincere, no matter how the rest of the group behaves.

In fact, it has been observed experimentally that in referendums imposing participation quorums, agents that expect to be in the minority often abstain (Aguiar-Conraria and Magalhães, 2010). This also relates to the well-known *no-show paradox* of voting theory (Fishburn and Brams, 1983), stressing that there are situations where agents may achieve a preferable outcome by not participating in the collective decision making process. Consider the following example.

5.2. Example. In continuation of Example 3.17, suppose that Isabelle (agent i), the president of the college, does not want to offer fresh fruit on campus. Suppose also that there exists a quorum of 21% that needs to be reached in order for the collective decision to have an effect: at least 21 out of the 100 members of the board must express *some* judgment on the fruit issue, and at least half of the reported judgments must be supportive, for the board to decide to offer the fruit. If 20 people wish to have the fruit and these are all the board members that actually have an opinion on the issue, then Isabelle has two options: to express her judgment and end up having to offer the fruit, or to insincerely abstain and achieve her preferred outcome. $\triangle$

Suppose more generally that agent i is sincerely in favour of some proposition $\varphi \in \Phi$. Then, depending on the context, agent i may try to force φ into the collective outcome, exclude $\neg\varphi$ from the collective outcome, or both.

5.3. Example. Imagine a local referendum within a municipality where, in case both φ and $\neg\varphi$ end up in the outcome, the mayor will make the final decision either by accepting φ or by accepting $\neg\varphi$, according to his own judgment. If agent i is optimistic and believes that the mayor will agree with her, then she will settle for having φ in the outcome and giving the mayor the option to choose it; if i is pessimistic and believes that the mayor—given the option—will explicitly disagree with her, then she will try to exclude $\neg\varphi$ from the outcome. $\triangle$

Below, we formalise the notion of manipulation in judgment aggregation with possibly incomplete inputs. We take into consideration all possible types of manipulation that may occur according to our discussion. Specifically, we make two main assumptions:

(i) An agent with a sincere positive judgment on $\widetilde{\varphi}$ will not seek the opportunity to manipulate against φ or favourably to $\neg\varphi$.[1]

(ii) An agent who sincerely does not hold an opinion on $\widetilde{\varphi}$ will not seek the opportunity to manipulate regarding $\widetilde{\varphi}$.[2]

Definition 5.4 concerns resolute aggregation rules, which return a single judgment set as the collective decision. For an example of such rules, the reader may keep in mind the quota rules that we presented in Section 3.2—in contrast, the weight rules of Section 3.1.1 are irresolute, possibly returning multiple judgment sets, and thus our definition does not directly apply to them.

5.4. Definition. *An agent i with sincere judgment J_i is said to have an* ***opportunity to manipulate*** *the aggregation rule F on proposition φ if $\varphi \in J_i$ and there exist a profile $\boldsymbol{J} = (\boldsymbol{J}_{-i}, J_i)$ and an insincere judgment J'_i such that the following holds:*

$$F(\boldsymbol{J}_{-i}, J_i) \lessdot_{\varphi} F(\boldsymbol{J}_{-i}, J'_i) \quad \text{or} \quad F(\boldsymbol{J}_{-i}, J'_i) \lessdot_{\neg\varphi} F(\boldsymbol{J}_{-i}, J_i)$$

An agent favouring φ may possibly add both φ and $\neg\varphi$ to the outcome. We consider such cases as opportunities for manipulation too. The aggregation rule F is *manipulable* if there exists an agent i with sincere judgment J_i that has an opportunity to manipulate F on some proposition $\varphi \in \Phi$. Otherwise F is *immune to manipulation*.

The notion of manipulation in judgment aggregation adopted in this section—in line with the original one by Dietrich and List (2007c)—does not involve any definition of individual *preferences*: A manipulation act by an agent may be triggered because of several behavioural reasons. Examples 5.2 and 5.3 show concrete contexts where the *opportunities* of agents to manipulate translate into *incentives* to manipulate.[3]

[1] This assumption may be violated in case an agent has a reason to want her sincere judgment collectively refuted. For this section, we exclude such special circumstances and assume that the sincere opinions of the agents about the issues at stake—in case they exist—overrule other possible motives.

[2] This assumption may also be violated for agents that sincerely have no opinion on $\widetilde{\varphi}$ but may still try to steer the collective outcome towards a positive or a negative decision on $\widetilde{\varphi}$. When accounting for this kind of manipulation, monotonicity ceases to be a sufficient property for a quota rule's immunity to manipulation. We discuss this in more detail towards the end of this section.

[3] Much of the recent work on strategic behaviour in judgment aggregation with complete inputs relies on preferences, like those defined based on the *Hamming distance*. The original characterisation result of Dietrich and List (2007c) does not hold for such classes of preferences, which are more restrictive. See Baumeister et al. (2017) for a survey on the topic.

Intuitively, the manipulation illustrated in Example 5.2 was possible due to a failure of monotonicity: Additional support on proposition φ was able to turn the collective outcome against φ. Next, we prove more generally that monotonicity together with weak independence are sufficient conditions for any aggregation rule to be immune to manipulation (Lemma 5.5). This means in particular that all monotonic quota rules for incomplete judgments are immune to manipulation.

5.5. Lemma. *Any aggregation rule for incomplete judgments that is judgment monotonic and weakly independent will be immune to manipulation.*

Proof. Aiming for a contradiction, consider an aggregation rule F under which some agent i has an opportunity to manipulate on a proposition φ with $\varphi \in J_i$. If F is weakly independent, i cannot unilaterally change the outcome on φ by keeping her judgment on $\widetilde{\varphi}$ the same. So, assume that i reports an insincere judgment J_i' with $J_i \neq_\varphi J_i'$ instead. This means that $J_i' <_\varphi J_i$. If F is judgment monotonic, it will hold that $F(\boldsymbol{J}_{-i}, J_i') \leq_\varphi F(\boldsymbol{J}_{-i}, J_i)$ and $F(\boldsymbol{J}_{-i}, J_i) \leq_{\neg\varphi} F(\boldsymbol{J}_{-i}, J_i')$. It follows from Definition 5.4 that i will not have an opportunity to manipulate on φ, which is a contradiction. □

We further show that judgment monotonicity is a necessary condition to be satisfied by any non-manipulable aggregation rule for incomplete judgments (see Lemma 5.6). So, we can identify exactly those quota rules for incomplete inputs that offer a barrier to manipulation (see Proposition 5.7).

5.6. Lemma. *Any aggregation rule for incomplete inputs that is immune to manipulation will satisfy judgment monotonicity.*

Proof. Consider an aggregation rule F that violates judgment monotonicity: There exist a proposition $\varphi \in \Phi$ and two profiles $\boldsymbol{J} = (\boldsymbol{J}_{-i}, J_i), \boldsymbol{J}' = (\boldsymbol{J}_{-i}, J_i')$ such that $J_i' <_\varphi J_i$ and either $F(\boldsymbol{J}_{-i}, J_i) <_\varphi F(\boldsymbol{J}_{-i}, J_i')$ or $F(\boldsymbol{J}_{-i}, J_i') \leq_{\neg\varphi} F(\boldsymbol{J}_{-i}, J_i)$. Clearly, an agent i with sincere judgment J_i will have an opportunity to manipulate F on proposition φ. □

Proposition 5.7 follows from the fact that the only quota rules satisfying judgment monotonicity are the invariable ones (consult Table 4.1 in Section 4.2.1).

5.7. Proposition. *A quota rule for incomplete judgments is immune to manipulation if and only if it satisfies judgment monotonicity, that is, if and only if it is an invariable (absolute or marginal) quota rule.*

> The violation of weak independence (as opposed to that of judgment monotonicity) does not directly imply the manipulability of an aggregation rule. To see this, suppose that an agent sincerely abstaining on $\widetilde{\varphi}$ unilaterally alters her judgment regarding a different issue $\widetilde{\psi}$, causing a change in the collective outcome on $\widetilde{\varphi}$. Although this is a violation of weak independence, the relevant definitions do not prescribe that

it constitutes an opportunity for manipulation on φ. If, nonetheless, we modify our definitions to incorporate the possible opportunities that agents who abstain on an issue may find to manipulate, then monotonicity will not anymore provide a safety net against the manipulability of a rule: Given a monotonic rule, an agent may strategically add a proposition φ in her judgment set in order to promote φ's acceptance in the outcome (or, arguably more plausibly, in order to harm the collective acceptance of $\neg\varphi$), even if the agent should sincerely abstain on $\widetilde{\varphi}$. This contrasts with the well-known result of Dietrich and List (2007c), who characterised all non-manipulable aggregation rules for the special case of complete inputs in terms of the two axioms of independence and monotonicity.

We close this section by providing a characterisation of all non-manipulable aggregation rules for incomplete judgments, which is aligned with the aforementioned result by Dietrich and List. To that end, we define an even weaker independence axiom that gives the freedom to agents that abstain on $\widetilde{\varphi}$ to still influence, indirectly, the collective decision on $\widetilde{\varphi}$. More precisely, *very weak independence* captures the idea that the collective outcome on a proposition φ should only depend on the judgments regarding $\widetilde{\varphi}$ of the agents that have an opinion directly on $\widetilde{\varphi}$ or possibly on the judgments regarding different issues $\widetilde{\psi}$ of the agents that abstain on $\widetilde{\varphi}$.

Axiom: ***Very Weak Independence.*** *For all propositions $\varphi \in \Phi$ and profiles $\boldsymbol{J} = (J_1, \dots, J_n)$, $\boldsymbol{J}' = (J'_1, \dots, J'_n)$, whenever $N^{\boldsymbol{J}}_{\varphi} = N^{\boldsymbol{J}'}_{\varphi}$, $N^{\boldsymbol{J}}_{\neg\varphi} = N^{\boldsymbol{J}'}_{\neg\varphi}$, and $J_i = J'_i$ for all $i \notin N^{\boldsymbol{J}}_{\widetilde{\varphi}}$, it holds that $F(\boldsymbol{J}) =_{\varphi} F(\boldsymbol{J}')$.*

5.8. Theorem. *An aggregation rule for incomplete judgments is immune to manipulation if and only if it satisfies judgment monotonicity and very weak independence.*

Proof. The "if" is proven analogously to Lemma 5.5. For the "only if", we know from Lemma 5.6 that any non-monotonic aggregation rule for incomplete inputs will be susceptible to manipulation. It remains to show that any aggregation rule for incomplete inputs that is not very weakly independent will be susceptible to manipulation.

Consider a rule F that violates very weak independence: There exist proposition $\varphi \in \Phi$ and profiles $\boldsymbol{J}, \boldsymbol{J}'$ such that $N^{\boldsymbol{J}}_{\varphi} = N^{\boldsymbol{J}'}_{\varphi}$, $N^{\boldsymbol{J}}_{\neg\varphi} = N^{\boldsymbol{J}'}_{\neg\varphi}$, and $J_i = J'_i$ for all $i \notin N^{\boldsymbol{J}}_{\widetilde{\varphi}}$, but $F(\boldsymbol{J}) \neq_{\varphi} F(\boldsymbol{J}')$. Suppose that $F(\boldsymbol{J}) <_{\varphi} F(\boldsymbol{J}')$ ($F(\boldsymbol{J}') <_{\varphi} F(\boldsymbol{J})$ is treated analogously).

We first define the profile $\boldsymbol{J}^{\mathbf{1}} = \boldsymbol{J} = (J_1, \dots, J_n)$. Then, for $j \in \{2, \dots, n+1\}$, we define the following profiles:

$$\boldsymbol{J}^{\boldsymbol{j}} = (J'_1, \dots, J'_{j-1}, J_j, \dots, J_n)$$

Since $F(\boldsymbol{J}) <_{\varphi} F(\boldsymbol{J}')$, we can identify an agent i who is the first one that increases the support of φ in the outcome by changing her judgment. Formally, the following holds:

$$i \in \operatorname*{argmin}_{j \in \{1, \dots, n\}} \boldsymbol{J}^{\boldsymbol{j}} <_{\varphi} \boldsymbol{J}^{\boldsymbol{j+1}}$$

In the two profiles $\boldsymbol{J}^{i}$ and $\boldsymbol{J}^{i+1}$ exactly the same agents support φ, exactly the same agents support $\neg\varphi$, and the agents that abstain on $\widetilde{\varphi}$ report exactly the same judgments. Then, $\boldsymbol{J}^{i} \neq \boldsymbol{J}^{i+1}$ implies that that $J_i \neq J'_i$, which means that agent i does not abstain on $\widetilde{\varphi}$. Now if $\varphi \in J_i$, agent i has an opportunity to manipulate by reporting J'_i instead of J_i; if $\neg\varphi \in J_i$ (which also means that $\neg\varphi \in J'_i$), she has an opportunity to manipulate when sincerely holding the judgment J'_i by reporting J_i instead. □

Having concluded our study within judgment aggregation, we will proceed with voting.

5.2 Voting: The One-Shot Case

Imagine that you take part in a voting process. The other members in your group have already submitted their preferences, and you actually have access to them (for example because you are the chair of a board, or because you know your friends—and their opinions—very well). After you also report your preference, a specific positional scoring rule is going to be applied to aggregate all given preferences, and some (possibly tied) alternative(s) will be selected as the winner(s). This will be an one-shot procedure, meaning that no preference will be subject to change after its submission. You are free to report an incomplete preference over the alternatives in question. Naturally, you may then consider what are the outcomes that you can induce by reporting different kinds of preferences: your sincere preference, a preference that is very different from your sincere one, or a preference that is just a slight modification of your sincere one. Your action will depend on the change you will be able to cause to the collective decision, as well as on the social and practical expectations regarding your behaviour (for example, in some circles not communicating a piece of information is not considered a real lie).

It is the fact that we allow for incomplete preferences that encourages novel types of manipulation. Besides *omitting* pairwise preferences, you may also invent preferences that you do not really hold, thereby *adding* pairwise preferences; you may also say you prefer one alternative to another when in fact the reverse is true, thereby *flipping* your preference on some pair of alternatives. For each of these types, and their combinations, we will provide necessary and sufficient conditions for positional scoring rules to make the relevant type of manipulation unprofitable.

> Positional scoring rules are exactly those voting rules that satisfy the axioms of anonymity, neutrality, continuity, and reinforcement (recall Section 4.3.1). So, in all the results of this section, the reader could replace the term "positional scoring rule" with "any rule that satisfies the four relevant axioms".

We formalise three types of lying:

- asserting something that you believe to be false;
- refraining from asserting something that you believe to be true;

- asserting something where in reality you have no opinion.

See Example 5.9 for an illustration.

5.9. Example. Suppose that your sincere preference is as follows:

$$\begin{tikzcd} a \arrow[r] \arrow[rr, bend left] & b \arrow[r] & c \end{tikzcd}$$

Further suppose that the domination scoring rule of Section 3.3 is used: We count how often an alternative is preferred over other alternatives, and select the alternatives with the highest count. Then, it may be profitable to omit pairwise preferences where b is preferred, submitting the following insincere preference:

$$\begin{tikzcd} a \arrow[r] \arrow[rr, bend left] & b & c \end{tikzcd}$$

Because this decreases the count of b, the new preference is a better one to submit in order to make a win (and since your sincere preference ranks a above b, we assume that you prefer that a wins rather than b).[4] △

The manipulation move of Example 5.9 would not be admissible if the agents were required to submit complete preferences.

In certain situations one may be more worried about one type of manipulation than another. For instance, in crowdsourcing settings, part of the mechanism designer's problem is not only to get the agents to tell the truth, but also to incentivise them to fully complete the task. Lying by omission may deprive the system of important data. On the other hand, in social situations, lying by omission intuitively seems more acceptable than outright misrepresentation. So, some manipulation moves may be considered tolerable. But this may also lead to an inverse conclusion for a mechanism designer: If concealing a fragment of their truth feels more acceptable to agents in a particular scenario than creating a brand new lie or completely twisting their preferences, it may be thought as more important to protect against the more likely type of lie.

The literature on manipulation in voting was pioneered by Gibbard (1973) and Satterthwaite (1975) who independently proved that, given some fairly weak conditions, no voting rule is immune to strategic manipulation. The conditions include requiring that the agents report complete preferences, as well as that the voting rule is resolute (i.e., always outputs a single winner).

Later work has relaxed the single-winner condition, which leads to the question of how to extend preferences over alternatives to preferences over sets of alternatives (still for complete preferences). Early contributions concerning this issue were made by

[4] In general, preferring alternative a over alternative b and preferring that a wins rather than that b wins should be distinguished. For example, you may deem a beach better than all other beaches in a competition without wanting it to win the award, because you think it will then be overrun by tourists. Here, we only consider situations where the two kinds of preferences can be deemed the same.

Gärdenfors (1976) and Kelly (1977), followed more recently by Ching and Zhou (2002) and Sato (2008). Specifically, two widely studied preference extensions are those of *optimistic* and of *pessimistic* voters. The former prefer a set of alternatives X over a set of alternatives Y if and only if their favourite alternative in X is better than their favourite alternative in Y; the latter prefer X over Y if and only if their worst alternative in X is better than their worst alternative in Y. We know by Duggan and Schwartz (2000) that any reasonable[5] voting rule is immune to manipulation by both optimistic and pessimistic voters if and only if it is *weakly dictatorial* (where a weak dictatorship for complete preference profiles is an irresolute voting rule in which there exists a certain agent that always has her *unique* top alternative in the collective outcome).

Pini et al. (2008) provided another version of Duggan and Schwartz's result, considering agents with incomplete preferences (that can be misrepresented in every possible way), and showing that it is still impossible to avoid weak dictators (where a weak dictator for incomplete preference profiles is an agent who always has *some* of her top alternatives in the outcome). Building on this stream of research, this section focuses on specific types of manipulation of practical interest and brings out some good news. Note that if we allow the agents to lie in any way they want, then only *trivial* positional scoring rules constitute weak dictatorships amongst all positional scoring rules—by restricting attention to concrete the manipulation types, we will find that some non-trivial positional scoring rules become immune to manipulation as well.

> All our results are independent of the specific preference domain (i.e., transitive or not), as long as we work with strict preferences. For ease of exposition, here we will work with strict and transitive preferences $\rhd \in \mathcal{P}(A)^t$. Moreover, we will assume an infinite superpopulation throughout the section, and no additional assumption will be made regarding the set of all potential alternatives.

5.2.1 Definitions

The notion of a desirable manipulation move is fairly obvious when there are only singleton outputs. Then, one can directly see from an agent's preference if she will prefer a different singleton output obtained through manipulation.

A preference $\rhd_i$ indicates which single alternatives are preferred to other single alternatives by agent i. As a starting point, we assume the following:

$$\text{if } a \rhd_i b, \text{ then agent } i \text{ will prefer } \{a\} \text{ over } \{b\}$$

This sets a minimal condition on an agent's preference. But positional scoring rules generally output non-singleton sets of alternatives. Hence, we need to determine when agent i wants to manipulate from an arbitrary set X to an arbitrary set Y, for $X, Y \subseteq A$.

[5] Where "reasonable" means that it gives the opportunity to every alternative to be elected as the winner under some preference profile (i.e., it satisfies nonimposition).

There is a broad literature concerning elaborate extensions of preferences from singletons to preferences over sets of objects (Barberà et al. (2004) provide a survey). Typically, extensions aim to describe the induced preferences of agents over sets of objects, given various assumptions about the agent's psychology (e.g., being risk-adverse or risk-loving) and about how the alternative sets will actually be consumed (e.g., a random alternative must be used or all alternatives must be used). Our intended interpretation of extensions is directly connected to the notion of manipulation. More specifically, a preference extension in our model does not need to represent the preferences of some particular agent, but rather the possible preferences that someone might conceivably have. The idea is that if *some* alternative in a new collective outcome is preferred to *some* alternative in the old outcome, then an agent may desire to manipulate from the old to the new. In fact, we will cover even more cases, allowing one or other of the alternatives to be contained in both sets. Definition 5.10 introduces our notion of preference extension via two requirements.

5.10. Definition. *Given a preference $\rhd$, a binary relation $\rhd^\star$ over sets of alternatives is called a **$\star$-extension** of $\rhd$ if and only if the following two conditions hold:*

(*i*) *$x \rhd y$ implies that $\{x\} \rhd^\star \{y\}$, for all $x, y \in A$*

(*ii*) *$X \rhd^\star Y$ implies that there exist $x \in X$ and $y \in Y$ such that $x \rhd y$ and $\{x, y\} \not\subseteq X \cap Y$, for all $X, Y \subseteq A$*

Note that the largest binary relation that can be called an extension according to Definition 5.10 may not be acyclic; it may even contain symmetries.[6] This underlines the fact that a preference extension in our terminology is not intended to be a particular representation of preferences over sets of objects. We do not intend to mean that for a given agent the set X is preferred to Y *and* that Y is preferred to X, but rather that it is conceivable that X is preferred to Y *or* that Y is preferred to X. Preventing all manipulation moves induced by any kind of preference extension would constitute a strong positive result. Moreover, most known definitions of concrete preference extensions in the literature, like the optimistic and the pessimistic ones previously sketched, constitute extensions in the line of Definition 5.10.

5.11. Example. Suppose that we have sincere preferences as follows.

Ann

$$\begin{tikzcd} a \arrow[r] \arrow[d] & b \arrow[dl] \\ c & d \end{tikzcd}$$

Bea

$$\begin{tikzcd} a & b \arrow[l] \arrow[d] \\ c & d \end{tikzcd}$$

Cal

$$\begin{tikzcd} a & b \\ c \arrow[u] \arrow[r] & d \end{tikzcd}$$

For the domination-scoring rule, alternatives a and c have total score 2, while alternative d has score 0 and alternative b is the winner with total score 3. But Ann can make

[6]So, it is possible that both $X \rhd^\star Y$ and $Y \rhd^\star X$.

$\{a, c\}$ be the collective outcome instead: by omitting the pairwise comparison (b, c) and adding (a, d), the score of a will increase to 3 and the score of b will decrease to 2. According to the largest possible preference extension $\star$, this constitutes a manipulation move that may be desirable by an agent in Ann's position with preference $\rhd$, since it holds that $\{a, c\} \rhd^\star \{b\}$. Note that it is also the case that $\{b\} \rhd^\star \{a, c\}$. $\triangle$

We have described when an outcome is more desirable for a manipulator; we still need to determine who the manipulators are, and how they can achieve this outcome.

Classic studies of manipulation consider a single agent acting alone. However, many modern applications of preference aggregation require a broader definition of manipulative acts. An Internet user can now easily create and handle several identities (all sharing the same sincere preference).[7] Our model allows for an—arbitrarily large—group of agents with identical preferences to manipulate.[8]

5.12. Definition. *Consider a voting rule F, an extension $\star$, a subset of agents $N' \subseteq N$ with the preference $\rhd$, and the following profile:*

$$\blacktriangleright = (\rhd_1, \ldots, \rhd_n) = (\underbrace{\rhd, \ldots, \rhd}_{\ell}, \blacktriangleright_{-N'}) \in \mathcal{P}(A)^n$$

*Without loss of generality, suppose that $N' = \{1, \ldots, \ell\}$. A **$\star$-manipulation** of F by N' is possible if there exists a new profile $(\rhd'_1, \ldots, \rhd'_\ell)$ such that the following folds:*

$$F(\rhd'_1, \ldots, \rhd'_\ell, \blacktriangleright_{-N'}) \rhd^\star F(\blacktriangleright)$$

The rule F is *manipulable* if a manipulation is possible by some subgroup $N' \subseteq N$, for some N; F is *immune to manipulation* otherwise.

Definition 5.12 tells us who can manipulate. It does not tell us in what manner. We sum up the three types of misrepresentation that are relevant for incomplete preferences with reference to a typical commitment of sworn testimony:

> One tells the truth, the whole truth, and nothing but the truth.

This is decomposed into three parts, prohibiting three different types of manipulation.[9] Consider a $\star$-manipulation of F by N' (Definition 5.12). Such a manipulation is by:

- *flipping* if $\rhd_i \backslash \rhd'_i = \{(x, y) \mid (y, x) \in \rhd'_i \backslash \rhd_i\}$ for all $i \in N'$;

[7] The idea of group-manipulation is also related to the notion of *false-name* manipulation (Yokoo et al., 2001) and to the so-called *Sybil attacks* (Douceur, 2002).

[8] Besides its practical significance, allowing for group-manipulation gives stronger sufficient conditions for preventing manipulation. In the other direction, this assumption does not weaken much our necessity resluts, since we only use at most two agents for the proofs, and the second agent may be considered as a somewhat technical requirement in order to break ties.

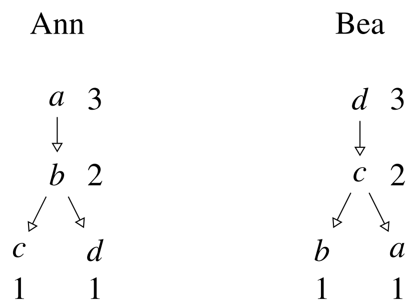
[9] We assume that all manipulators share the same manipulation type. This is justified in scenarios where behind all the identities of the manipulators there is one single agent,

- *omission* if $\rhd_i \supseteq \rhd'_i$ for all $i \in N'$;
- *addition* if $\rhd_i \subseteq \rhd'_i$ for all $i \in N'$.

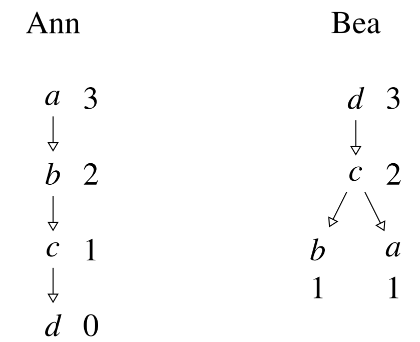
We can also naturally define manipulation by *combinations* of the above types. For instance, a manipulation is by a *combination of addition and omission* if the agents add some pairwise preferences and possibly omit some others. We say that two preferences $\rhd$ and $\rhd'$ *conform to* a certain (combination of) manipulation type(s) if $\rhd$ and $\rhd'$ satisfy the corresponding relations.

To further illuminate manipulation moves of different types, let us recall the rules in the Borda class of Section 3.3.2, for top-truncated preferences.

5.13. Example. Consider a profile of sincere preferences as follows (where transitive arrows are not depicted), and suppose that we employ the optimistic Borda rule.



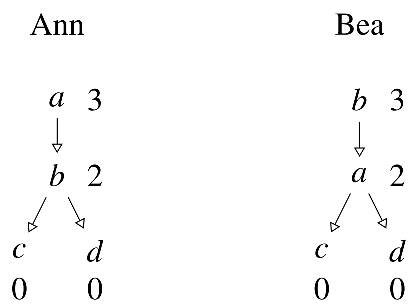
Then, the winning set is $\{a, d\}$. But Ann prefers alternative a over d, so she could manipulate in order to have the singleton $\{a\}$ as the outcome:



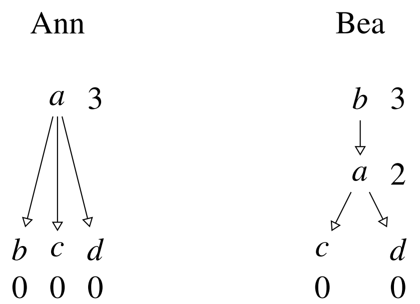
The depicted manipulation move of Ann incorporates the addition of a pairwise preference from alternative c to alternative d. △

The manipulation move of Example 5.13 would not work if the pessimistic Borda rule was applied instead. However, both the optimistic and the pessimistic Borda rules are manipulable by omission, as Example 5.14 shows.

5.14. Example. Consider a profile of sincere preferences as follows (where transitive arrows are not depicted), and suppose that we employ the pessimistic Borda rule.



Then, the winning set is $\{a, b\}$. But Ann prefers alternative a over b, so she could manipulate in order to have the singleton $\{a\}$ as the outcome:



Ann incorporated omission of pairwise preferences from alternative b to alternatives c and d. The same move would be effective under the optimistic Borda rule too. △

5.2.2 Conditions for Immunity to Manipulation

Certain positional scoring rules can be immune to manipulation for all possible preference extensions (Theorem 5.16). This follows from our sufficient conditions (which, as we will see, are not always easy to satisfy). In the other direction, we present necessary conditions for a rule to be immune to manipulation for at least one (the smallest) extension, thus establishing vital boundaries for non-manipulability (Theorem 5.17). Lemma 5.15 is critical for these results.

5.15. Lemma. *If a positional scoring function s is not trivial, then for all positive numbers $t > 0$ and all alternatives x and y, there exists a profile $\rhd$ such that for all $z \neq x, y$ the following holds:*

$$\sum_{i \in N} s_{\rhd_i}(x) = \sum_{i \in N} s_{\rhd_i}(y) > t + \sum_{i \in N} s_{\rhd_i}(z)$$

Proof. Take two arbitrary alternatives $x, y \in A$. For a non-trivial scoring function s there is a preference $\rhd$ for which we can order the alternatives using indices $j \in \{1, \ldots, |A|\}$ such that $s(a_j) \geqslant s(a_{j+1})$ and $s(a_1) > s(a_{|A|})$. Consider a permutation of this preference where x is placed in the position of a_1, y is placed in the position of a_2, some $z \neq x, y$ is placed in the position of a_3, etc. Now iteratively create new preferences which swap the positions of x and y and cycle through the positions of the other alternatives: After $2|A| - 4$ iterations this results in a collection of preferences for which x is in the position of a_1 exactly $|A| - 2$ times and in the position of a_2 exactly $|A| - 2$ times; where y is in those positions exactly as often as x; and where, for each $j = \{3, \ldots, |A|\}$, every other alternative is in the position a_j exactly twice. Obviously x and y have the same summed score for these preferences. Similarly, all other alternatives have the same summed scores as each other. The inequalities over the scores of ordered alternatives imply that the summed score of x, and that of y, is some value $\delta > 0$ larger than the summed score of any other alternative. The required profile is created by taking $1 + \lceil t/\delta \rceil$ copies of the $2|A| - 4$ preferences. □

Given preferences $\rhd$ and $\rhd'$ and alternatives $x, y \in A$, let us define the following inequality, stating that the difference between the scores of x and y in $\rhd'$ is not strictly larger than the relevant difference in $\rhd$:

$$s_{\rhd}(x) - s_{\rhd}(y) \geqslant s_{\rhd'}(x) - s_{\rhd'}(y) \tag{5.1}$$

Thanks to Theorem 5.16 below, we will be able to make conclusions such as the following: A positional scoring rule is immune to manipulation by omission if it guarantees that inequality (5.1) holds for every two alternatives that are ranked by $\rhd$ and for every preference $\rhd'$ that is obtained trough $\rhd$ by omission.

5.16. Theorem. *If inequality (5.1) holds for all preferences $\rhd, \rhd'$ and all alternatives $x, y \in A$ such that $x \rhd y$ and $\rhd, \rhd'$ conform to a specific (combination of) type(s), then the rule F_s, induced by the positional scoring function s, is immune to $\star$-manipulation by the given type(s), for all extensions $\star$.*

Proof. Aiming for a contradiction, suppose that the condition of the statement holds, but F_s is $\star$-manipulable by the given type(s). This means that there exist $X, Y \subseteq A$ such that the following holds:

$$X = F(\boldsymbol{\rhd}_{-N'}, \rhd'_1, \ldots, \rhd'_k) \rhd^\star F(\boldsymbol{\rhd}_{-N'}, \rhd, \ldots, \rhd) = Y$$

for some subset $N' = \{1, \ldots, k\} \subseteq N$ of a group N and insincere preferences $\rhd'_j$ such that $\rhd$ and $\rhd'_j$ conform to the given type(s) for all $j \in \{1, \ldots, k\}$. Then, by the definition of preference extensions, there are $x \in X, y \in Y$ such that $x \rhd y$ and $\{x, y\} \nsubseteq X \cap Y$. We focus on the case where $x \notin Y$, since the case where $y \notin X$ is analogous. Because $y \in Y$, by the definition of the scoring rule we have that the following is the case:

$$\sum_{i \in N'} s_{\rhd}(x) + \sum_{i \in N \setminus N'} s_{\rhd_i}(x) < \sum_{i \in N'} s_{\rhd}(y) + \sum_{i \in N \setminus N'} s_{\rhd_i}(y)$$

Then, $x \in X$ implies the following:

$$\sum_{i \in N'} s_{\rhd'_i}(x) + \sum_{i \in N \setminus N'} s_{\rhd_i}(x) \geqslant \sum_{i \in N'} s_{\rhd'_i}(y) + \sum_{i \in N \setminus N'} s_{\rhd_i}(y)$$

We arrive to a final inequality:

$$\sum_{i \in N'} (s_{\rhd_i}(x) - s_{\rhd_i}(y)) < \sum_{i \in N'} (s_{\rhd'_i}(x) - s_{\rhd'_i}(y))$$

Thus, we have contradicted our hypothesis. □

Since avoiding manipulation by a group of agents is harder than avoiding manipulation by a single agent, we know that the condition of Theorem 5.16 is sufficient to avoid manipulation by just a single agent as well.

We continue with our necessary conditions. Theorem 5.17 enables us to make conclusions such as the following: A positional scoring rule that is immune to manipulation by omission must guarantee that inequality (5.1) holds for every two alternatives that are ranked by $\rhd$ and for every preference $\rhd'$ that is obtained trough $\rhd$ by omission.

5.17. Theorem. *For all extensions $\star$, if the rule F_s, induced by the positional scoring function s, is immune to $\star$-manipulation by a specific (combination of) type(s), then inequality (5.1) holds for all $\rhd, \rhd' \in \mathcal{P}(A)$ and $x, y \in A$ such that $\rhd$ and $\rhd'$ conform to the given type(s) and $x \rhd y$.*

Proof. We prove the contrapositive. Suppose that there exist preferences $\rhd, \rhd'$ and alternatives x, y as in the statement that satisfy $s_\rhd(x) - s_\rhd(y) < s_{\rhd'}(x) - s_{\rhd'}(y)$. This implies that the scoring function is not trivial.

Consider the following profile with four agents:

$$\boldsymbol{\rhd}' = (\rhd, \rhd, \rhd_{\langle xy \rangle}, \rhd'_{\langle xy \rangle})$$

From the inequality and because s is positional, y must have a higher score than x in the profile $\boldsymbol{\rhd}'$. However, if the first two agents change their preference to $\rhd'$, then x has a higher score than y. We can write this profile as $\boldsymbol{\rhd}'' = (\rhd', \rhd', \rhd_{\langle xy \rangle}, \rhd'_{\langle xy \rangle})$.

Let the following to equalities hold:

$$t = \max_{z \in A} \sum_{i \in N} s_{\rhd_i}(z) - \sum_{i \in N} s_{\rhd_i}(y)$$
$$t' = \max_{z \in A} \sum_{i \in N} s_{\rhd'_i}(z) - \sum_{i \in N} s_{\rhd'_i}(x)$$

We use Lemma 5.15 to create a profile $\boldsymbol{\rhd}'''$ in which the scores of x and y are equal and greater than the score of any other alternative by at least $\max(t, t')$.

The profile *from* which the manipulation happens is $(\boldsymbol{\rhd}', \boldsymbol{\rhd}''')$, where y is winning; the profile *to* which the manipulation happens is $(\boldsymbol{\rhd}'', \boldsymbol{\rhd}''')$, where x is winning. □

Theorems 5.16 and 5.17 imply a characterisation result.

5.18. Corollary. *For all extensions $\star$, the rule F_s, induced by the positional scoring function s, is immune to $\star$-manipulation by a specific (combination of) type(s) if and only if inequality (5.1) holds for all $\triangleright, \triangleright' \in \mathcal{P}(A)$ and all $x, y \in A$ such that $\triangleright$ and $\triangleright'$ conform to the given type and $x \triangleright y$.*

Our characterisation result facilitates the detection of whether a particular positional scoring rule is immune to manipulation, focusing on pairs of preferences instead of full preference profiles. Nonetheless, we do not know whether it will be computationally easy to check that the required condition is satisfied by a positional scoring rule.

5.2.3 Manipulation by Specific Types

We have so far formulated general conditions for immunity to manipulation. We can now proceed to examine specific cases regarding our types of interest.

We will initially study whether it is possible for a non-trivial positional scoring rule to prevent combined types of manipulation. The answer we obtain is largely negative, but is also contingent to each particular manipulation type. Then, we will continue with analysing each single manipulation type in its own right.

Combined Types

Our first remark concerns the combined type of addition and flipping, and is promising. The stepwise scoring rule (defined in Section 3.3.1, page 45) satisfies the condition of Theorem 5.16, instantiated for preferences $\triangleright$ and $\triangleright'$ that conform to flipping: Take two such preferences $\triangleright$ and $\triangleright'$ and two alternatives a and b such that $a \triangleright b$. If $b \triangleright' a$, then the difference between the stepwise scores of a and b cannot increase—it will either remain the same, or decrease. If $a \triangleright' b$, then the relevant difference will remain exactly the same. Analogous reasoning applies for addition, and straightforwardly for the combination of addition and flipping, leading to Proposition 5.19.

5.19. Proposition. *There exists a non-trivial positional scoring rule that is immune to $\star$-manipulation by the combination of addition and flipping, for all extensions $\star$.*

However, we obtain an impossibility result regarding immunity to manipulation by the combination of omission and flipping. Theorem 5.23 below is implied by Lemma 5.21 and Lemma 5.22. Lemma 5.20 is needed for the proof of the latter (and follows from Theorem 5.17 instantiated for omission, together with the fact that all positional scoring functions assign the same score to all alternatives in the totally empty preference).

5.20. Lemma. *For all extensions $\star$, if the rule F_s, induced by the positional scoring function s, is immune to $\star$-manipulation by omission, then $s_\triangleright(x) \geqslant s_\triangleright(y)$ for all preferences $\triangleright$ and alternatives x, y with $x \triangleright y$.*

5.21. Lemma. *For all extensions $\star$, if the rule F_s, induced by the non-trivial positional scoring function s, is immune to $\star$-manipulation by omission, then for all complete preferences $\triangleright$ there exist $x, y \in A$ such that $s_{\triangleright}(x) \neq s_{\triangleright}(y)$.*

Proof. Consider a positional scoring function s and a rule F_s immune to $\star$-manipulation by omission. In every empty preference, all alternatives get the same score by s (condition $(*)$). Suppose that all alternatives in every complete preference also get the same score (condition $(**)$). We will show that all alternatives must have the same score in every preference $\triangleright$, and thus s must be trivial.

Take an arbitrary preference $\triangleright$ and alternatives x, y with $x \triangleright y$. Theorem 5.17 instantiated for omission and applied on $(*)$ and $(**)$ implies that $s_{\triangleright}(x) = s_{\triangleright}(y)$. Thus, all connected alternatives will have the same score (condition$(* * *)$).

But all alternatives x, y that are not connected must also be assigned with the same score: We can create two new preferences by connecting x and y by (i) an arrow from x to y and (ii) an arrow from y to x (each of these two arrows can be added without creating a cycle, since x and y were not connected in the first place). Theorem 5.17 instantiated for omission and applied on (i), (ii) and $(* * *)$ implies that the relevant scores must be the same, and the proof is concluded. □

5.22. Lemma. *For all extensions $\star$, if the rule F_s, induced by the positional scoring function s, is immune to $\star$-manipulation by the combination of omission and flipping, then for all complete preferences $\triangleright$ and alternatives x, y, it holds that $s_{\triangleright}(x) = s_{\triangleright}(y)$.*

Proof. Consider an arbitrary complete preference $\triangleright$ for which it holds that $s_{\triangleright}(x) \geqslant s_{\triangleright}(y)$ whenever $x \triangleright y$ (this must be the case by Lemma 5.20, if the rule F_s is immune to manipulation by omission).

For some ordering of the alternatives $\{a_1 \ldots, a_m\} = A$, and without drawing the transitive arrows for simplicity, the preference $\triangleright$ will look as follows, where the scores assigned to the alternatives are mentioned below them (for $\gamma \in \mathbb{R}$):

$$\begin{array}{ccccccccc} a_1 & \longrightarrow & a_2 & \longrightarrow & a_3 & \longrightarrow & \cdots & \longrightarrow & a_m \\ \gamma & & \gamma - \delta_1 & & \gamma - \delta_1 - \delta_2 & & \cdots & & \gamma - \sum_{i=1}^{m-1} \delta_i \end{array}$$

We know that $\delta_i \geqslant 0$, and will show that $\delta_i = 0$, for all $1 \leqslant i \leqslant m - 1$.

Consider flipping the preference between a_1 and a_2 as follows:

$$\begin{array}{ccccccccc} a_2 & \longrightarrow & a_1 & \longrightarrow & a_3 & \longrightarrow & \cdots & \longrightarrow & a_m \\ \gamma & & \gamma - \delta_1 & & \gamma - \delta_1 - \delta_2 & & \cdots & & \gamma - \sum_{i=1}^{m-1} \delta_i \end{array}$$

Because the difference in scores between a_2 and a_3 cannot increase (by Theorem 5.17 instantiated for flipping), we must have that $\delta_1 = 0$.

Now flip the preference between a_1 and a_3:

$$\begin{array}{ccccccccc} a_2 & \longrightarrow & a_3 & \longrightarrow & a_1 & \longrightarrow & \cdots & \longrightarrow & a_m \\ \gamma & & \gamma & & \gamma - \delta_2 & & \cdots & & \gamma - \sum_{i=2}^{m-1} \delta_i \end{array}$$

In order for the difference in scores between a_2 and a_1 to not increase (again by Theorem 5.17), it must hold that $\delta_2 = 0$.

We repeat this process, "moving" a_1 towards the bottom of the preference in steps, and obtaining $\delta_i = 0$ for all $1 \leqslant i \leqslant m$. □

5.23. Theorem. *For all extensions $\star$, only the trivial positional scoring rule is immune to $\star$-manipulation by the combination of omission and flipping.*

Next, we directly prove an impossibility result regarding immunity to manipulation by the combination of addition and omission.

5.24. Theorem. *For all extensions $\star$, only the trivial positional scoring rule is immune to $\star$-manipulation by the combination of addition and omission.*

Proof. Consider a rule F_s, induced by a positional scoring function s, that is immune to $\star$-manipulation both by addition and by omission. We will prove that for any preference $\rhd$, $s_\rhd(x) = s_\rhd(y)$ for all $x, y \in A$. If $\rhd$ is empty, we are done. So suppose there is a pair $(a, b) \in \rhd$. For each such pair (x, y) in $\rhd$, define the (singleton) preference $\rhd_{xy} = \{(x, y)\}$. Theorem 5.17 for omission and for addition implies that for all (x, y) in $\rhd$, the following holds:

$$s_\rhd(x) - s_\rhd(y) = s_{\rhd_{xy}}(x) - s_{\rhd_{xy}}(y) \tag{5.2}$$

Since the scoring rules we consider are positional, by Equation (5.2) the following holds for all $(x, y), (x', y')$ in $\rhd$:

$$s_\rhd(x) - s_\rhd(y) = s_\rhd(x') - s_\rhd(y') \tag{5.3}$$

Consider two alternatives a and b such that $a \rhd b$ and define the preference $\rhd' = \{(a, b), (b, c), (a, c)\}$ for some $c \in A \setminus \{a, b\}$. Again by Theorem 5.17 for addition and omission, $s_{\rhd_{ab}}(a) - s_{\rhd_{ab}}(b) = s_{\rhd'}(a) - s_{\rhd'}(b)$. By Equation (5.3) applied to the pairs (a, b), (b, c), and (a, c), it follows that $s_{\rhd'}(a) = s_{\rhd'}(b) = s_{\rhd'}(c)$.

We thus know that $s_\rhd(x) = s_\rhd(y)$ whenever x and y are connected. Suppose now that x and y belong to two different connected components of $\rhd$. Then, the preferences $\rhd' = \rhd \cup \{(x, y)\}$ and $\rhd'' = \rhd \cup \{(y, x)\}$ will be acyclic and thus well-defined. By the same reasoning as above, we know that $s_{\rhd'}(x) = s_{\rhd'}(y)$ and $s_{\rhd''}(x) = s_{\rhd''}(y)$. Since $\rhd \subset \rhd'$ and $\rhd \subset \rhd''$, it follows by Theorem 5.17 for omission that $s_\rhd(x) - s_\rhd(y) \leqslant s_{\rhd'}(x) - s_{\rhd'}(y) = 0$ and that $s_\rhd(y) - s_\rhd(x) \leqslant s_{\rhd''}(y) - s_{\rhd''}(x) = 0$. So, $s_\rhd(x) = s_\rhd(y)$ and we have concluded the proof. □

Theorems 5.23 and 5.24 are strong, in the sense that their proofs do not require that specific manipulation moves employ both omission and flipping (or addition and omission) at the same time—they only require that an aggregation rule be immune to all manipulation moves that employ some single type from the relevant ones.

Single Types

After having investigated manipulation by combined types and obtained two impossibility results, we now wonder whether specific types of manipulation (by addition, omission, or flipping) can at least be prevented by positional scoring rules. We discover very positive facts: For every single manipulation type, there exists some positional scoring rule immune to manipulation.

First, recall that the stepwise scoring rule prevents manipulation by both addition and flipping (Proposition 5.19). So, we immediately know the following:

5.25. Proposition. *There exists a non-trivial positional scoring rule that is immune to $\star$-manipulation by addition, for all extensions $\star$.*

5.26. Proposition. *There exists a non-trivial positional scoring rule that is immune to $\star$-manipulation by flipping, for all extensions $\star$.*

Note though that the stepwise scoring rule must be manipulable by omission, otherwise the impossibilities for the combined types would fail. See Example 5.27.

5.27. Example. Consider the stepwise scoring function, and the following preference:

$$\begin{tikzcd} a \arrow[r] \arrow[d] \arrow[dr] & c \\ b \arrow[r] & d \end{tikzcd}$$

Omitting the arrow from b to d increases the difference in score between a and b. △

We also find that 1-veto scoring satisfies the condition of Theorem 5.16, instantiated for preferences that conform to addition. Indeed, if $a \rhd b$ for some alternatives a, b, then the veto score of a must be 1, the 1-veto score of b must be 0, and the two scores will remain the same under every addition of pairwise preferences to $\rhd$. But 1-veto is manipulable by omission and by flipping (Example 5.28).

5.28. Example. Consider the preference below.

$$\begin{tikzcd} \mathrm{a} \arrow[r] \arrow[rr, bend left] & \mathrm{b} \arrow[r] & \mathrm{c} \end{tikzcd}$$

Then, we can increase the difference between the 1-veto scores of a and b either by omitting or by flipping the arrow from b to c. △

Next, Proposition 5.29 holds because the cumulative scoring rule satisfies the condition of Theorem 5.16, instantiated for preferences $\rhd$ and $\rhd'$ that conform to omission: If a is preferred to b with respect to preference $\rhd$, then by removing pairwise preferences from $\rhd$ the cumulative score of a will be reduced at least as much as that of b.

5.29. Proposition. *There exists a non-trivial positional scoring rule that is immune to $\star$-manipulation by omission, for all extensions $\star$.*

But omission is the only manipulation type to which the cumulative scoring rule is immune. Example 5.30 illustrates this.

5.30. Example. In the left preference below, we can increase the difference between the cumulative scores of a and b by adding arrows from a to c and from a to d.

$$a \rightarrow b \qquad c \rightarrow d \qquad\qquad \begin{tikzcd} a \arrow[r] & b \arrow[r] & c \end{tikzcd}$$

In the right preference, we can flip the arrow connecting a and b to increase the difference between b and c. △

Figure 5.1 graphically depicts our observations. The existence of rules that are immune to manipulation by flipping and not by addition remains a conjecture.

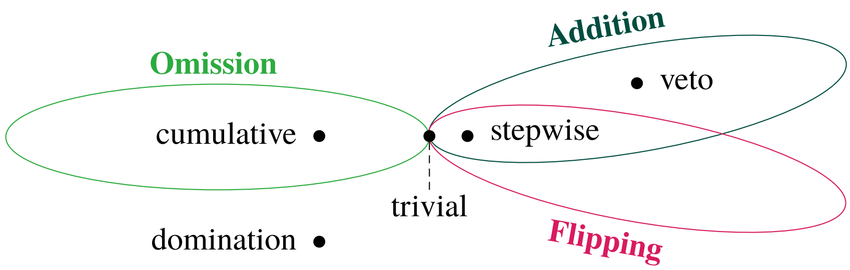


Figure 5.1: Some positional scoring rules, categorised with respect to their immunity to manipulation by the different types of omission, addition, and flipping.

To sum up, in the social choice literature scoring rules are thought of as largely manipulable. Here, we validate this idea. Positional scoring rules are very much "all or nothing": Either they prevent a huge amount of manipulation moves (for all possible preference extensions) or they are clearly manipulable (again, for all preference extensions, and using just two identical agents). It is an open question whether our impossibilities can be circumvented outside the family of positional scoring rules, and how those rules that are immune to manipulation can be characterised axiomatically.

5.3 Voting: The Iterative Case

Iterative voting features in settings where a collective decision is spread over multiple rounds (Meir, 2017). A common example is the `Doodle` tool, allowing its users to modify their reported preferences over time in light of new information they have possibly acquired. The analysis of iterative voting within computational social choice was initiated a decade ago by Meir et al. (2010) and has been a hot topic since then; researchers have been especially concerned with those iterative processes that are based on scoring rules, such as plurality and antiplurality (Reyhani and Wilson, 2012; Brânzei et al., 2013; Lev and Rosenschein, 2016; Obraztsova et al., 2015a; Meir et al., 2017).[10] Importantly, the literature to date in the area has been relying directly on the restrictive assumption that all agents hold complete preferences.

It thus seems essential to consider voting models accommodating iteration and incomplete preferences at the same time, and this is the topic of this section.[11]

We study agents that are *strategic*, meaning that after observing the submissions of their peers, they modify their own preferences—one-by-one—in order to obtain a better outcome. In line with previous literature (Lev and Rosenschein, 2016; Meir et al., 2017), our agents are *myopic*: they only attempt an improved outcome in the immediate next round, without thinking further ahead. The main question we are interested in concerns the problem of convergence, and is summarised as follows:

> Do iterative voting processes with incomplete preferences reach stable states, where no agent has an incentive (or is able) to change her preference?

Convergence properties are crucial for the design and effective use of collective decision making tools: In their absence, it may be impossible to reach a final decision.

We recall two ways of generalising the k-approval rules for standard voting with incomplete preferences (including the plurality and antiplurality rules) from Section 3.3.1, namely the k-approval family and the k-veto family. For the purposes of this section, we will ensure a unique winning alternative in every round of the iterative process. In order to do so, we pair the voting rules with a lexicographic tie-breaking rule that, given an outcome $S \subseteq A$, selects the alternative a when $a \in S$, the alternative b when $b \in S$ and $a \notin S$, and so on. This pairing will always be implied in this section.

Aside from analysing the general convergence qualities of rules and comparing our results to existing ones from the literature, we also investigate the impact of specific types of preference change on convergence. We find that no rule is guaranteed to always converge. This finding contrasts with seminal convergence results by Meir et al. (2010), Lev and Rosenschein (2012), and Reyhani and Wilson (2012) regarding plurality and antiplurality under linear preferences. Yet, all rules can attain a convergence assurance under certain restrictions, which we characterise. Whenever convergence is achieved, we evaluate the maximum number of rounds required for a stable state to emerge.

In the classical framework of iterative voting with linear orders, various other restrictions have been imposed to facilitate convergence results, typically concerning behavioural characteristics of the agents like *truth-biasedness*, *farsightedness*, *laziness*, and *incomplete information* (Reijngoud and Endriss, 2012; Grandi et al., 2013; Obraztsova et al., 2013; Meir et al., 2014; Rabinovich et al., 2015; Endriss et al., 2016).

---

[10] Exceptions exist; for instance, Endriss et al. (2016) and Koolyk et al. (2017) have also studied voting rules that are not based on scoring functions.

[11] Dery et al. (2019) have also explored iterative voting under incompleteness. However, their work assumes that it is not the actual preferences of the agents, but rather the mechanism designer's information about these preferences, that is incomplete.

These restrictions differ in nature from the ones we consider, as they cannot be easily put in place by an external mechanism designer. Obraztsova et al. (2015b) have provided a general account on how severe such restrictions actually need to be.

To illustrate the open questions that iterative voting with incomplete preferences brings to the surface, let us see an example.

5.31. Example. Consider Ann and Bea, who hold the sincere preferences $a \rhd b \rhd c \rhd d$ and $d \rhd b \rhd c \rhd a$, respectively. Suppose that we employ the 2-approval rule, and the agents are only allowed to report complete preferences. In the initial sincere profile, alternative b is the winner. However, Ann prefers alternative a over b, which she can make the winner by changing her preference to $a \rhd c \rhd b \rhd d$ with her first manipulation move. Then, Bea can make alternative c win with another manipulation move (the second one), since she prefers c over a. Ann subsequently has an incentive to return to her sincere preference in order to make a win (that she prefers over c), which will be followed by the manipulation move of Bea, also returning to her sincere preference to make b win. This process (depicted in Figure 5.2) would lead to a cycle, with manipulation moves possibly being repeated *ad infinitum*.

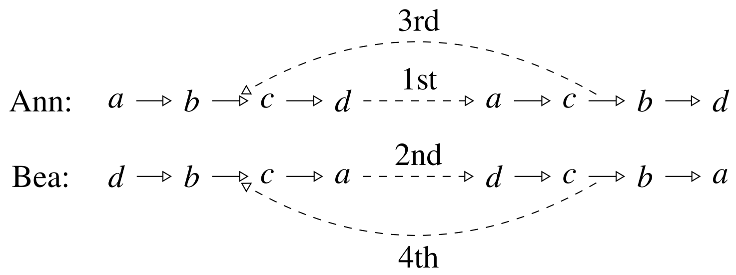


Figure 5.2: A cycle when only complete preferences are allowed.

Suppose that Ann and Bea are granted the freedom to report incomplete preferences, while they still hold the same (complete) preferences as above. In the initial sincere profile, alternative b is winning under the 2-approval rule. Then, Ann can make her most preferred alternative a win be placing b in a lower position, which she does in the first manipulation round. Now, Bea needs to increase the score of alternative c (that she prefers over a), by placing it below at most one other alternative. She has two ways for doing that: either by flipping her pairwise preference between b and c, or by simply omitting that preference. Suppose that she chooses the latter option. Ann can subsequently report her sincere preference to make alternative b the winner (that she prefers over c). In the end, Bea does not have the possibility to improve the outcome for herself after Ann's last move, so the process will reach a stable state (Figure 5.3).

How agents change their preferences is thus crucial for convergence. $\triangle$

Let us now set the stage for our formal results, by establishing the appropriate definitions. To start, a process of iterative voting based on a voting rule F takes place in rounds.

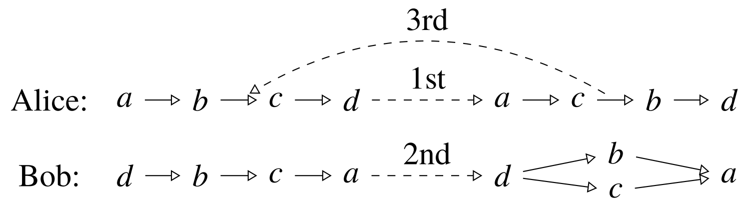


Figure 5.3: Convergence when incomplete preferences are allowed.

Given a scoring function s and two alternatives $x, y \in A$, we write $s_t(x) <_\ell s_t(y)$ when either (i) the score of x is strictly smaller than the score of y in round t, or (ii) the score of x is the same as the score of y in round t, but the lexicographic tie-breaking rule ranks y above x. For simplicity, we then say that the score of x is smaller than the score of y, without referring to the tie-breaking rule.

The winner in every round is determined through F. We assume the following:

- ◇ In the very beginning (round 1), all agents are sincere. (*assumption 1*)

Assumption (1) is reasonable, since agents that have no information about the preferences of their peers have no incentive to submit an insincere preference (Reijngoud and Endriss, 2012). In each subsequent round, an agent may change her reported preference only if this will induce a better outcome for herself. Concretely, denoting by $\blacktriangleright^t_{-i}$ the profile of the rest of the group in round t, agent i with sincere preference $\rhd_i$ may change her reported preference $\rhd^t \in \mathcal{P}(A)$ to $\rhd^{t+1} \in \mathcal{P}(A)$ only if the following holds:

$$F(\blacktriangleright^t_{-i}, \rhd^{t+1}) \rhd_i F(\blacktriangleright^t_{-i}, \rhd^t)$$

We also make the following assumption:

- ◇ When more than one agent can profit by changing their reported preference, the agent that does so first will be the one that has her move registered. We impose no further restriction on who this agent might be. (*assumption* 2)

Since we generally allow for incomplete preferences, an agent i may modify her reported preference $\rhd^t_i \in \mathcal{P}(A)$ to a new one $\rhd^{t+1}_i \in \mathcal{P}(A)$ in three fundamental ways, or any combination thereof, as defined in Section 5.2: by omission, addition, or flipping.

The iterative process based on a voting rule F converges in profile $\blacktriangleright$ when no agent has an incentive (or is able) to further change her preference in $\blacktriangleright$. The profile $\blacktriangleright$ is then called a *stable* state.[12] We also say that the rule F is *guaranteed convergence* if it always reaches a stable state, starting from any sincere profile, for every number of agents n from an infinite superpopulation, and of alternatives m from an infinite set of

[12]Stable states do not always constitute Nash equilibria in the classical game-theoretical sense, since agents may still want to unilaterally deviate from them, without being allowed to do so.

potential alternatives. Assumption (2) implies that when we say that an iterative process is guaranteed to reach a stable state, this will include all possible orders in which the agents might act. Straightforwardly, for voting rules that are immune to manipulation in an one-shot procedure (recall Section 5.2), convergence is achieved in 0 rounds.

We study agents choosing to report preferences that constitute *best responses*: If agent i, who sincerely prefers a to b, can make either a or b the winner by reporting a new preference, then she will choose to make her preferred alternative a win.[13] After an agent has identified a best response, we additionally assume the following:

- ⋄ The agent will not alter her reported preference more than necessary to achieve her desirable outcome. (*assumption 3*)

We say that the agents' moves are *minimally diverging*. Assumption (3), based on the idea that changing one's preference is costly, is necessary (yet not always sufficient) for our convergence results. Specifically, flipping one's preference on a pair of alternatives will be taken to cost more than just removing or just adding the relevant pairwise comparison. This can be externally imposed in application contexts. For example, in an online poll the designer can force the participants to delete an already submitted ranking between two alternatives before they are able to report a different ranking regarding the same alternatives.[14] Formally, for two preferences $\rhd, \rhd'$ and a pair of alternatives $(x, y) \in A^2$, we define $C_{xy}(\rhd, \rhd')$ to be the *cost* of moving between the two preferences with respect to the given pair (and we say that the two preferences *fully agree* on (x, y) if they both rank x above (or below) y, or if they both deem x incomparable to y; and they *directly disagree* on (x, y) if one of them ranks x above y, while the other one ranks y above x):

$$C_{xy}(\rhd, \rhd') = \begin{cases} 0 & \text{if } \rhd, \rhd' \text{ fully agree on } (x, y) \\ 1 & \text{if } \rhd, \rhd' \text{ fully disagree on } (x, y) \\ 1/2 & \text{otherwise} \end{cases}$$

The *total cost* of moving from $\rhd$ to $\rhd'$ is defined as follows:

$$C(\rhd, \rhd') = \sum_{(x,y)\in A^2} C_{xy}(\rhd, \rhd')$$

If agent i with preference $\rhd_i^t$ in round t can make a preferable alternative x win by reporting either preference $\rhd$ or preference $\rhd'$, and if it also holds that $C(\rhd_i^t, \rhd) < C(\rhd_i^t, \rhd')$, then agent i will report preference $\rhd$.

Lastly, consider agent i with sincere preference $\rhd_i$, who modifies her reported preference in round t in order to change the winner from w_t to w_{t+1}. We catergorise the possible moves of the agent into two types, that are not necessarily disjoint:

[13] Best responses do not need to be unique; e.g., two incomparable alternatives may simultaneously constitute best responses. In case of multiple best responses, the agents randomly select one of them.

[14] Agents may also find flipping more costly because lying in that fashion intuitively is more extreme.

- *Type 1*: $s_{t+1}(w_t) <_\ell s_t(w_t)$. That is, agent i decreases the score of w_t to make some other alternative the winner;

- *Type 2*: $s_t(w_{t+1}) <_\ell s_{t+1}(w_{t+1})$. That is, agent i increases the score of w_{t+1} to make it the winner.

> This section is based on possibly incomplete preferences that are strict and transitive. Transitivity in particular is an essential assumption that we take on board here. If we drop it, then several of our convergence results will cease to hold. To understand why this might be important, consider for instance an agent who submits the preference $\{ab, bc, ac\}$. If the agent needs to omit the pairwise comparison ac from her preference, then she will have to omit some other pair as well in order to keep transitivity in place (thus increasing the cost of her action).

In the proofs of Propositions 5.32–5.52 that will be presented later on, given a profile $\blacktriangleright$, we will capture the scores of the m alternatives in A (in alphabetical order) with a vector in $\mathbb{R}^m$, depicting the winning score in boldface.

5.3.1 The Approval Family

This section explores convergence for the approval family.

1-approval

We know that the plurality rule—the counterpart of 1-approval for linear orders—is guaranteed to converge when the agents are only allowed to report complete preferences (Meir et al., 2010). This result does not survive in richer settings of incompleteness. In fact, there exists a counterexample for convergence of the iterative 1-approval process, where the agents only alter a single pairwise comparison in their sincere preferences.

5.32. Proposition. *If omission and addition are allowed, then the iterative 1-approval process is not guaranteed to converge (independently of whether flipping is allowed).*

Proof. Consider a sincere preference profile as follows:

Agent 1 : $c \rhd d \rhd a \rhd b$	Agent 5 : $d \rhd c \rhd b \rhd a$
Agent 2 : $b \rhd d \rhd c \rhd a$	Agent 6 : $d \rhd b \rhd c \rhd a$
Agent 3 : $a \rhd d \rhd b \rhd c$	Agent 7 : $c \rhd b \rhd d \rhd a$
Agent 4 : $b \rhd c \rhd d \rhd a$	Agent 8 : $a \rhd b \rhd c \rhd d$

The first seven agents will form a cycle; every time one of them makes a move, she will be altering the position of the alternative ranked second in her sincere preference.

$$
\begin{tikzcd}
(\mathbf{2},2,2,2) \arrow[r, "1d\uparrow"'] & (2,2,2,\mathbf{3}) \arrow[r, "4c\uparrow"'] & (2,2,\mathbf{3},3) \arrow[r, "2d\uparrow"'] & (2,2,3,\mathbf{4}) \arrow[d, "1d\downarrow"] \\
(2,\mathbf{3},3,3) \arrow[d, "3d\uparrow"] & (2,3,\mathbf{4},3) \arrow[l, "4c\downarrow"'] & (2,\mathbf{3},3,3) \arrow[l, "5c\uparrow"'] & (2,2,\mathbf{3},3) \arrow[l, "6b\uparrow"'] \\
(2,3,3,\mathbf{4}) \arrow[r, "7b\uparrow"] & (2,\mathbf{4},3,4) \arrow[r, "6b\downarrow"] & (2,3,3,\mathbf{4}) \arrow[r, "2d\downarrow"] & (2,\mathbf{3},3,3) \arrow[dll, "7b\downarrow"] \\
(2,2,2,\mathbf{3}) \arrow[uuu, bend left, "3d\downarrow"] & (2,2,\mathbf{3},3) \arrow[l, "5c\downarrow"'] & &
\end{tikzcd}
$$

For example, consider agent 1 with sincere preference $c \rhd d \rhd a \rhd b$: We write "$1d\uparrow$" when agent 1 moves alternative d on top by omission to make d win instead of a, and "$1d\downarrow$" when she removes alternative d from the top by addition and submits her sincere preference to make c win, as illustrated below:

$$
\begin{tikzcd}
 & & & & c \arrow[dr] & & \\
c \arrow[r] & d \arrow[r] & a \arrow[r] & b \arrow[r, dashed, "1d\uparrow"] & {} & a \arrow[r] \arrow[lllll, bend right=40, dashed, "1d\downarrow"', start anchor=north east] & b \\
 & & & & d \arrow[ur] & &
\end{tikzcd}
$$

In the sincere preference of agent 1, alternative c gets 1 point and all other alternatives get 0 points; in the manipulated preference, alternatives c and d get 1 point each and all other alternatives still get 0 points. □

We can, though, obtain convergence if we restrict the ways the agents are allowed to change their preferences. Our proofs (and certain proofs in later parts of this paper too) are akin to those of Reyhani and Wilson (2012) for the plurality and antiplurality rules under linear preferences, showing that the set of potentially winning alternatives decreases along a path of best responses. Some of the ideas we develop—for example the fact that after a flipping move the non-winning alternative loses its chance to become a winner in the future—also featured in the original work of Meir et al. (2010).

Some of our results in this section will be stated for $(m-1)$-veto in parallel with 1-approval (and will be called upon later on, in the veto section), whenever the reasoning for the two rules is analogous.

Proposition 5.34 will rely on the assumptions of transitivity and of minimal divergence, while this will not be the case for Proposition 5.35.

> Both in Proposition 5.34 and in Proposition 5.35, we prove that the agents will only perform moves of type 2, and convergence is a logical consequence of this (in the former case, because omission moves can be done finitely many times; in

the latter case, because making a new alternative the winner can be done finitely many times). But the underlying reasons for our proofs' validity are very different (linked to either the absence of the *ability* or the absence of the *interest* to make the relevant manipulation moves, respectively). On the one hand, when addition is not allowed, the agents would only want to perform a move of type 1 by flipping, but all relevant moves would violate transitivity and would thus not be admissible. On the other hand, when omission is not allowed, the agents' incentives will always concern moves of type 2, and their desired actions will only employ flipping.

We will often use Lemma 5.33, which is a consequence of the definitions of our rules.

5.33. Lemma. *The first manipulation move of every agent under the iterative 1-approval process or the iterative* $(m-1)$*-veto process will be of type 2.*

5.34. Proposition. *If addition is not allowed, 1-approval is guaranteed to converge.*

Proof. We will show by induction that if addition is not allowed, then an arbitrary agent i with sincere preference $\rhd_i$ will also never employ flipping. She will be changing her submitted preference only by omission, making a new alternative the winner each time. Since this can happen for at most finitely many rounds, convergence follows.

Our induction basis is established by Lemma 5.33 (a move of type 2 will be done by omission, due to the minimal divergence assumption). Suppose now that the statement holds for the k^{th} time agent i changes her preference. We will show that it also holds for the $(k+1)^{\text{th}}$ time she changes her preference. Say this happens in round t.

Case 1: Agent i's move in round t is of type 2 (and not of type 1). Then, omission will be employed as for round 1, and a new alternative will be placed on top.

Case 2: Agent i's move in round t is of type 1, preventing an alternative y from winning instead of a superior alternative x. Since addition is not allowed, the only way that y can have its score decreased is when y is a top alternative in $\rhd_i^t$ and agent i flips some pairwise preference $y \rhd_i^t z$ to $z \rhd_i^{t+1} y$. Note that $x \rhd_i y$ and $y \rhd_i z$ (because up until round t the agent has only omitted pairwise preferences), and $x \rhd_i z$ by transitivity. Since in round t alternative z is on top for agent i, $x \rhd_i^t z$ holds because the induction hypothesis tells us that only pairwise comparisons that aim at making a specific alternative on their dominated side the winner have been omitted. Hence, $x \rhd_i^{t+1} z$ also holds because the agents perform minimal changes on their submitted preferences. But then, $x \rhd_i^{t+1} z$ and $z \rhd_i^{t+1} y$ imply that $x \rhd_i^{t+1} y$ must be the case, by transitivity. This is impossible, since y is a top alternative in round t and addition is not allowed. □

5.35. Proposition. *If omission is not allowed, then the iterative 1-approval process and the iterative (m-1)-veto process are guaranteed to converge.*

Proof. We will prove the following claim: Every time an agent changes her preference, she promotes a winning alternative that is different from those she has promoted before. Because the set of alternatives A is finite, convergence follows.

Suppose that agent i is the first to change her reported (sincere) preference $\rhd_i$ in round 1. By Lemma 5.33, her move will be of type 2, aiming at increasing the score of some alternative y. For the 1-approval rule, this move will be done by flipping: Agent i will obtain a new pairwise preference with $y \rhd_i^2 x$, where x was a top alternative in $\rhd_i$ that was not winning. For the $(m-1)$-veto rule, the move may also be done by flipping (which will concern alternatives x and y as described for 1-approval), or it may be done by addition (which means that y was a top alternative in $\rhd_i$). We show the following:

(i) Every alternative that agent i sincerely places above y has no chance of winning in later rounds unless the score of the winner decreases.

We also know that the following holds:

(ii) If no better alternative can win, agent i will have no incentive to use a type-1 move to withdraw y from the winning position.

Condition (i) is clearly true for alternative x (if such x exists), which loses one point after i's move; it is also true for every non-top alternative z such that $z \rhd_i y$ (if such z exists), since z could not become a winner in round 2 (we know that because agent i submitted a best response), and the score of z remained the same after i's move. Lastly, suppose that there exists a top alternative r such that $r \rhd_i y$ and not $y \rhd_i^2 r$. As omission is not allowed, it should also hold that $r \rhd_i^2 y$, which is impossible since our rules must assign y score 1 in round 2.

Now, as is the case for agent i, every agent's first move will be of type 2. So the winning score will keep increasing until someone—say agent j—modifies her reported preference for the second time. Suppose that agent j made alternative u win with her first move. Because of (ii), agent j will also employ flipping for her second move that will be of type 2, making a new alternative $q \neq u$ win.

Continuing with this reasoning, we see that each agent who performs a second move will promote another alternative than the one she promoted in her first move; the same will be true for everyone who performs a third move, and so on. □

5.36. Corollary. *The iterative 1-approval process is guaranteed to converge if and only if omission is not allowed or addition is not allowed.*

The proofs of Proposition 5.34 and Proposition 5.35 show that when omission (but not addition) is allowed or when addition and flipping (but not omission) are allowed, a stable state will be reached by the iterative 1-approval process after at most nm rounds: Every time an agent changes her preferences, she moves a new alternative on top to make it win. When only addition is allowed, every sincere profile is a stable state.

k-approval, for $1 < k < m - 1$

By adjusting the non-convergence example for the 1-approval process under omission and addition, we construct one for k-approval, where $1 < k < m - 1$.

5.37. Proposition. *If omission and addition are allowed, then the iterative k-approval process is not guaranteed to converge (independently of whether flipping is allowed).*

Proof. Fix some $k \in \{2, \ldots, m-2\}$, and consider alternatives $e_1, \ldots, e_{k-1}$. In the profile provided in the proof of Proposition 5.32, place all new alternatives immediately after the top alternative of each agent. For example, for $k = 3$ the preference of agent 1 becomes $c \rhd e_1 \rhd e_2 \rhd d \rhd a \rhd b$. If needed, add some more agents to guarantee that the new alternatives have lower score than all other alternatives in the profile. □

Moreover, if we forbid omission, we find a non-convergence example for the iterative k-approval process where flipping is employed in literature (specifically, in the proof of Theorem 4, Part 1 of Lev and Rosenschein, 2016). Proposition 5.38 is thus true.

5.38. Proposition. *If flipping but not omission is allowed, then the iterative k-approval process is not guaranteed to converge (independently of whether addition is allowed).*

The conditions under which convergence of k-approval is guaranteed are more elaborate than those related to 1-approval. Particularly, in Proposition 5.39 we need to precisely describe the combinations of all allowed types of moves, as opposed to simply excluding addition (Proposition 5.34) or omission (Proposition 5.35).

5.39. Proposition. *If omission and flipping are allowed, but addition is not, then the iterative k-approval process is guaranteed to converge, for all* $1 < k < m - 1$.

Proof. Suppose that a cycle exists, aiming for a contradiction. The agents that participate in it will not employ omission, since such moves cannot be repeated when addition is not allowed. Then, all moves in the cycle will employ flipping. Moreover, all moves in the cycle will be of type 1: Indeed, flipping in some round t of the cycle will decrease the score of some alternative a in order to make another alternative win. If $a \neq w_t$, then the new winner could simply be obtained by omission (because of minimally diverging moves). So $a = w_t$, which is the definition of type 1.

Thus, in every round of the cycle the score of the winning alternative strictly decreases, and some alternative has its score increased by one point. Moreover, there must exist at least one alternative y that has its score increased an infinite number of times (since the cycle includes infinitely many steps, but the alternatives are finitely many). Then, alternative y must also have its score decreased an infinite number of times (otherwise it would be impossible for the score of the winner to always decrease). More specifically, every round t of the cycle where y has its score decreased must be followed by some round $t' > t$ where y has its score increased.

Consider such rounds t and t', where y does not have its score altered in rounds between t and t'. For all $x \in A \setminus \{y\}$, the following two conditions hold:

$$s_{t+1}(x) \leqslant_\ell s_{t+1}(w_{t+1}) <_\ell s_t(w_t)$$

$$s_t(w_t) = s_t(y) = s_{t+1}(y) + 1$$

Alternative y may have its score increased in round t' only if it is not a winner in t'. Then, we have the following:

$$s_{t'}(x) \leqslant_\ell s_{t'}(w_{t'}) \leqslant s_t(w_{t+1}) <_\ell s_t(w_t) = s_{t'}(y) + 1$$

Also, it holds that $s_{t'+1}(x) \leqslant_\ell s_{t'}(x)$, for all $x \in A \setminus \{y\}$. This means that if y has its score increased in round t' (that is, $s_{t'+1}(y) = s_{t'}(y) + 1$), y will be the winner in round $t' + 1$; but the agent can achieve this simply by increasing the score of y by one point, without changing the score of other alternatives. Given minimally diverging moves, this should happen by omission, which implies a contradiction. □

Keeping in mind that if only omission or only addition is allowed, then convergence follows straightforwardly, we can state Corollary 5.40.

5.40. Corollary. *The iterative k-approval process is guaranteed to converge, for all $1 < k < m - 1$, if and only if omission and addition are not allowed simultaneously, and flipping is allowed only when omission is allowed as well.*

Interestingly, under the assumption of minimally diverging moves, allowing more freedom to the agents—that is, omission on top of flipping—favours convergence.

When flipping and omission (but not addition) are allowed, iterative k-approval will reach a stable state after at most n^2m^2 rounds: Every agent performs omission at most m times, each of which with the purpose of increasing the score of a new alternative; between every two of these nm steps, at most nm flipping rounds materialise, where the score of the winner decreases and the score of some non-winner increases.

When only omission or addition is allowed, convergence occurs after at most nm rounds: Every agent's move consists of changing the score of a new alternative.

$(m-1)$-approval

Probably not surprisingly, we also find a counterexample for convergence also concerning the $(m - 1)$-approval process.

5.41. Proposition. *If flipping but not omission is allowed $(m - 1)$-approval is not guaranteed to converge (independently of whether addition is allowed as well).*

Proof. Consider a sincere preference profile as follows:

Agent 1 : $c \rhd a \rhd b \rhd d \rhd e$
Agent 2 : $d \rhd c \rhd b \rhd a \rhd e$
Agent 3 : $a \rhd d \rhd b \rhd c \rhd e$
Agent 4 : $b \rhd a \rhd c \rhd d \rhd e$
Agent 5 : $b \rhd d \rhd a \rhd c \rhd e$
Agent 6 : $b \rhd c \rhd d \rhd a \rhd e$
Agent 7 : $a \rhd c \rhd d \rhd e \rhd b$

The first six agents are going to participate in a cycle, starting from the upper left corner below; every time one of them makes a move, she will be altering the position of the alternative ranked second in her sincere preference.

$$\begin{tikzcd}
(\mathbf{7},6,7,7,1) \arrow[r, "1a\downarrow"'] & (6,6,\mathbf{7},7,2) \arrow[d, "2c\downarrow"'] & (5,\mathbf{6},6,6,5) \arrow[r, "3d\uparrow"'] & (5,6,6,\mathbf{7},4) \arrow[d, "5d\downarrow"'] \\
(6,6,\mathbf{7},7,2) \arrow[u, "4a\uparrow"'] & (6,6,6,\mathbf{7},3) \arrow[r, "3d\downarrow"] & (\mathbf{6},6,6,6,4) \arrow[u, "4a\downarrow"'] & (5,\mathbf{6},6,6,5) \arrow[d, "2c\uparrow"'] \\
(6,6,6,\mathbf{7},3) \arrow[u, "6c\uparrow"'] & (\mathbf{6},6,6,6,4) \arrow[l, "5d\uparrow"'] & (5,\mathbf{6},6,6,5) \arrow[l, "1a\uparrow"'] & (5,6,\mathbf{7},6,4) \arrow[l, "6c\downarrow"']
\end{tikzcd}$$

For example, consider agent 1 with sincere preference $c \rhd a \rhd b \rhd d \rhd e$. The first time that agent 1 performs a manipulation act, we write "$1a\downarrow$" when she moves alternative a at the bottom by flipping and submits the (minimally diverging) preference $c \rhd b \rhd d \rhd e \rhd a$ to prevent a be the winner instead of c. Then, we write "$1a\uparrow$" when she removes alternative a from the bottom to make a win instead of b, submitting the preference $c \rhd b \rhd d \rhd a \rhd e$. Finally, when agent 1 needs to decrease the score of a again (for which we also write "$1a\downarrow$"), she will submit the preference $c \rhd b \rhd d \rhd e \rhd a$. In the cycle, she will keep alternating between the two insincere preferences (see below).

$$\begin{tikzcd}
c \arrow[r] & a \arrow[r] & b \arrow[r] & d \arrow[r] & e \arrow[r, dashed, "1a\downarrow"] & c \arrow[r] & b \arrow[r] & d \arrow[r] \arrow[d, dashed, bend left, "1a\uparrow"] & e \arrow[r] & a \\
& & & & & c \arrow[r] & b \arrow[r] & d \arrow[r] \arrow[u, dashed, bend left, "1a\downarrow"] & a \arrow[r] & e
\end{tikzcd}$$

For all preferences of agent 1, all alternatives get 1 point each except for the bottom alternative that gets 0 points. □

> The non-convergence result for $(m-1)$-approval when only flipping is allowed (even when the agents all hold linear orders) contrasts with the fact that the antiplurality rule in standard voting always converges (Lev and Rosenschein, 2016). This happens because in the latter framework it is assumed that when an agent changes her reported preference, she always ranks the current winner in the lowermost position, while this is not the case under our minimal divergence assumption.

It is not difficult to convince ourselves that when all three types of omission, addition, and flipping are allowed, the counterexample in the proof of Proposition 5.41 continues to constitute a non-convergence case (where, for example, agent 1 will first employ flipping and then will alternate between increasing the score of a by omitting the pair (e, a) and decreasing the score of a by addition).

Once more, we see that when addition is not permitted, then allowing for omission on top of flipping is key to convergence (Proposition 5.43).

5.42. Proposition. *If omission and flipping are allowed, but addition is not, then the iterative* $(m - 1)$*-approval process is guaranteed to converge.*

Proof. The proof is identical to that of Proposition 5.39, regarding k-approval. □

When flipping and omission are allowed, but addition is not, then convergence will be achieved after at most n^2m^2 rounds, as we argued is the case for k-approval.

Furthermore, it is easy to see that when omission and addition are both allowed but flipping is not, convergence will be reached in at most n rounds: Every agent may only add a pairwise preference once, to reduce the score of an alternative by placing it at the bottom of a linear order (and when only omission or only addition is allowed, convergence will be reached after at most 0 and n rounds, respectively).

5.43. Proposition. *If omission and addition are allowed, but flipping is not, then the iterative* $(m - 1)$*-approval process is guaranteed to converge.*

Combining our observations, we now know what the conditions characterising convergence for the $(m - 1)$-approval rule are.

5.44. Corollary. *The iterative* $(m - 1)$*-approval process is guaranteed to converge if and only if flipping is not allowed, unless omission but not addition is allowed as well.*

5.3.2 The Veto Family

This section examines convergence for the veto family.

1-veto

For the 1-veto rule, we find a similar (in terms of the score-changes involved) counterexample as for $(m - 1)$-approval. However, different types of manipulation moves play a role now (specifically, omission and addition instead of flipping).

5.45. Proposition. *If omission and addition are allowed, then the iterative 1-veto process is not guaranteed to converge (independently of whether flipping is allowed).*

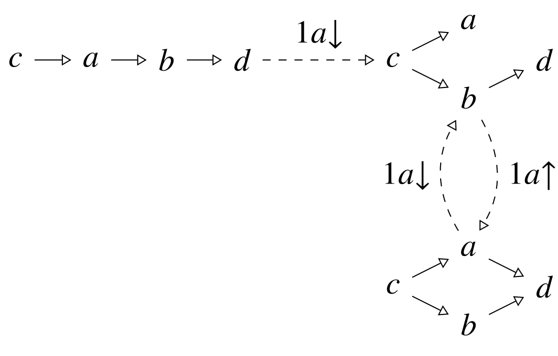
Proof. Consider a sincere preference profile as follows:

Agent 1 : $c \rhd a \rhd b \rhd d$	Agent 6 : $b \rhd c \rhd d \rhd a$
Agent 2 : $d \rhd c \rhd b \rhd a$	Agent 7 : $a \rhd c \rhd d \rhd b$
Agent 3 : $a \rhd d \rhd b \rhd c$	Agent 8 : $a \rhd c \rhd d \rhd b$
Agent 4 : $b \rhd a \rhd c \rhd d$	Agent 9 : $a \rhd c \rhd d \rhd b$
Agent 5 : $b \rhd d \rhd a \rhd c$	

The first six agents are going to participate in a cycle, starting at the upper left corner below; every time one of them makes a move, she alters the position of the alternative ranked second in her sincere preference.

$$\begin{tikzcd}
(\mathbf{7},6,7,7) \arrow[r, "1a\downarrow"'] & (6,6,\mathbf{7},7) \arrow[d, "2c\downarrow"'] & (5,\mathbf{6},6,6) \arrow[r, "3d\uparrow"'] & (5,6,6,\mathbf{7}) \arrow[d, "5d\downarrow"'] \\
(6,6,\mathbf{7},7) \arrow[u, "4a\uparrow"'] & (6,6,6,\mathbf{7}) \arrow[r, "3d\downarrow"] & (\mathbf{6},6,6,6) \arrow[u, "4a\downarrow"'] & (5,\mathbf{6},6,6) \arrow[d, "2c\uparrow"'] \\
(6,6,6,\mathbf{7}) \arrow[u, "6c\uparrow"'] & (\mathbf{6},6,6,6) \arrow[l, "5d\uparrow"'] & (5,\mathbf{6},6,6) \arrow[l, "1a\uparrow"'] & (5,6,\mathbf{7},6) \arrow[l, "6c\downarrow"']
\end{tikzcd}$$

For example, consider agent 1 with the sincere preference $c \rhd a \rhd b \rhd d$. The first time that agent 1 performs a manipulation act, we write "$1a\downarrow$" when she moves alternative a to the bottom by omission and submits the new preference $\{(c,a),(c,b),(b,d),(c,d)\}$, to make c win instead of a. Then, we write "$1a\uparrow$" when she removes a from the bottom by adding the pair (a,d), to make a win instead of b. When she needs to decrease the score of a again, agent 1 will just omit the pair (a,d) from her preference. In the cycle, she will alternate between the two insincere preferences. See below for an illustration.



In all preferences above, alternative d gets 0 points and alternatives c and b get 1 point each. What changes is that alternative a alternates between getting 1 and 0 points. □

When flipping but not omission is allowed, a similar counterexample to that in the proof of Proposition 5.45 can be constructed, in which only flipping is employed. That counterexample makes use of an additional alternative that will be moved back and forth from every agent's last position in the linear preference and will collect very low total score by all agents. Hence, we can provide Proposition 5.46.

5.46. Proposition. *If flipping but not omission is allowed, then the iterative 1-veto process is not guaranteed to converge (independently of whether addition is allowed).*

Again, we will study whether convergence can be guaranteed under further restrictions on the agents' moves.

5.47. Proposition. *If flipping and omission (but not addition) are allowed, then the iterative 1-veto process is guaranteed to convergence.*

Proof. Under the conditions of the statement, all moves will be of type 1 and will only use omission (the details simulate the proof of Proposition 5.34). □

Knowing that if only omission or only addition is allowed, then convergence will always be achieved, we obtain Corollary 5.48.

5.48. Corollary. *The iterative 1-veto process is guaranteed to converge if and only if omission and addition are not allowed simultaneously, and flipping is allowed only when omission is allowed as well.*

Convergence for 1-veto can be reached in at most nm rounds when flipping and omission (but not addition) are allowed or when only omission is allowed: Each agent will change her preference at most m times, to prevent a new alternative from being the winner. When only addition is allowed, all sincere profiles constitute stable states.

k-veto, for $1 < k < m - 1$

The analysis for k-veto draws inspiration from both our work regarding 1-veto and that regarding k-approval. First, all counterexamples for 1-veto can be reconstructed with some additional dummy alternatives, so that the relevant k comes into play.

5.49. Proposition. *If omission and addition are allowed simultaneously (independently of whether flipping is allowed), or if flipping is allowed without omission (independently of whether addition is allowed), then the iterative k-veto process is not guaranteed to converge.*

Concerning the positive results, if only omission or only addition is allowed, then convergence easily follows in at most mn steps. Moreover, we again find that allowing omission together with flipping is advantageous to convergence. The relevant proof is close to the proof of Proposition 5.39 regarding k-approval, although several of the technical tricks employed in the two cases are different—importantly, both proofs make explicit use of the minimal divergence assumption.

5.50. Proposition. *If flipping and omission (but not addition) are allowed, then the iterative k-veto process is guaranteed to converge.*

Proof. Suppose that a cycle exists, aiming for a contradiction. The agents that participate in it will not employ omission, since such moves cannot be repeated when addition is not allowed. Then, all moves in the cycle will employ flipping. Moreover, all moves in the cycle will be of type 2: Indeed, flipping in some round t of the cycle will increase the score of some alternative a and decrease the score of some alternative b, in order to make a the winner. If $a \neq w_{t+1}$, then the new winner could simply be obtained via decreasing the score of b by omission (because of minimally diverging moves). So $a = w_{t+1}$, which is the definition of type 2.

Thus, in every round of the cycle the score of the winning alternative strictly increases, and some alternative has its score decreased by one point. Moreover, in order for the score of the winner to always increase, there must exist at least one alternative y that has its score increased infinitely many times by a single agent i (since the cycle includes infinitely many steps, but the alternatives and the agents are finitely many). Then, every round t of the cycle where agent i decreases the score of y must be followed by some round $t' > t$ where agent i increases the score of y (in order to be possible for i to decrease and increase the score of y again).

Since alternative y has its score decreased in round t and the score of the winner strictly increases in the rounds following t, we know that y cannot become a winner after round t, including round $t' + 1$. Therefore, agent i would have no incentive to increase the score of y in round t', and we reached a contradiction. □

We know that in every sequence of flipping steps, all moves will be of type 2, constantly making new alternatives win. Such a sequence can take at most mn steps. Furthermore, an agent i may make use of omission in some round t in order to decrease the score of an alternative y and make a different alternative win. Suppose that there exists some alternative x such that $x \rhd_i^t y$ (if such an alternative does not exist, then it will apparently not be possible to change the score of y by flipping in the future). Note that for a decrease in the score of y to be realisable, there must also exist some suitable alternative z such that $y \rhd_i^t z$ (and thus $x \rhd_i^t z$ by transitivity). After the agent's omission move that will be minimally diverging, it will still be the case that $x \rhd_i^{t+1} y$ and $x \rhd_i^{t+1} z$, but not that $y \rhd_i^{t+1} z$. Since we assumed that addition is not allowed, it will not be possible for the agent to flip the pairwise preference between x and y to change the score of y in the future. Hence, the score of every alternative can be changed via omission at most n times. In total, when flipping and omission are allowed, convergence of the iterative k-veto process will be reached in at most $2mn$ steps.

5.51. Corollary. *The iterative k-veto process is guaranteed to converge, for all* $1 < k < m - 1$*, if and only if omission and addition are not allowed simultaneously, and flipping is allowed only when omission is allowed as well.*

($m - 1$)-veto

For $(m-1)$-veto, we identify a counterexample that in terms of scores is similar to the one for 1-approval, but in which the agents employ all three types of addition, omission, and flipping (as opposed to 1-approval, where only omission and addition were used).

5.52. Proposition. *When addition, omission, and flipping are all allowed, then the iterative $(m-1)$-veto process is not guaranteed to converge.*

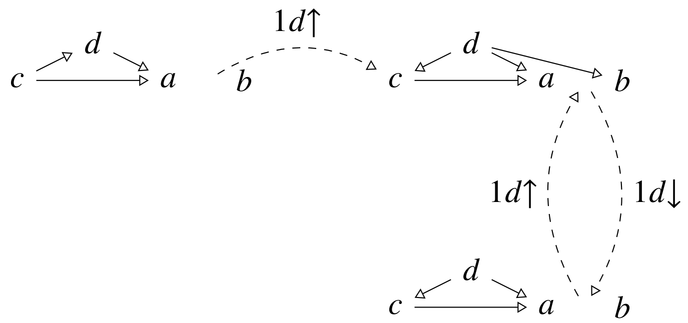
Proof. Consider a sincere preference profile as follows (where $(c \rhd d \rhd a), b$ abbreviates $\{(c,d),(c,a),(d,a)\}$):

$$
\begin{array}{ll}
\text{Agent } 1 : (c \rhd d \rhd a), b & \qquad \text{Agent } 5 : (d \rhd c \rhd b), a \\
\text{Agent } 2 : (b \rhd d \rhd c), a & \qquad \text{Agent } 6 : (d \rhd b \rhd c), a \\
\text{Agent } 3 : (a \rhd d \rhd b), c & \qquad \text{Agent } 7 : (c \rhd b \rhd d), a \\
\text{Agent } 4 : (b \rhd c \rhd d), a & \qquad \text{Agent } 8 : (a \rhd b \rhd c), d
\end{array}
$$

The first seven agents are going to participate in a cycle; every time one of them makes a move, she alters the position of the alternative ranked second in her sincere preference.

$$
\begin{tikzcd}
(\mathbf{0},0,0,0) \arrow[r, "1d\uparrow"'] & (0,0,0,\mathbf{1}) \arrow[r, "4c\uparrow"'] & (0,0,\mathbf{1},1) \arrow[r, "2d\uparrow"'] & (0,0,1,\mathbf{2}) \arrow[d, "1d\downarrow"] \\
(0,\mathbf{1},1,1) \arrow[d, "3d\uparrow"] & (0,1,\mathbf{2},1) \arrow[l, "4c\downarrow"'] & (0,\mathbf{1},1,1) \arrow[l, "5c\uparrow"'] & (0,0,\mathbf{1},1) \arrow[l, "6b\uparrow"'] \\
(0,1,1,\mathbf{2}) \arrow[r, "7b\uparrow"] & (0,\mathbf{2},1,2) \arrow[r, "6b\downarrow"] & (0,1,1,\mathbf{2}) \arrow[r, "2d\downarrow"] & (0,\mathbf{1},1,1) \arrow[dll, "7b\downarrow"] \\
(0,0,0,\mathbf{1}) \arrow[uuu, bend left, "3d\downarrow"] & (0,0,\mathbf{1},1) \arrow[l, "5c\downarrow"'] & &
\end{tikzcd}
$$

For example, consider agent 1 with sincere preference $(c \rhd d \rhd a), b$. In the first manipulation act of agent 1, we write "$1d\uparrow$" when she moves alternative d on the top by flipping and addition and submits the preference $\{(d,c),(c,a),(d,a),(d,b)\}$ to make d win instead of a. We write "$1d\downarrow$" when she removes alternative d from the top by omission and submits the preference $\{(d,c),(c,a),(d,a)\}$ to make c win. When agent 1 needs to increase the score of d again, she will simply add back the pairwise preference (d,b). In the cycle, agent 1 will alternate between the two insincere preferences, as depicted below.



When agent 1 places alternative d on top, then d gets score 1 and all other alternatives get score 0; in the other preferences of agent 1, all alternatives get score 0. □

Strikingly, the iterative $(m-1)$-veto process is not guaranteed convergence *only* if omission, addition, and flipping are all allowed.

5.53. Proposition. *If flipping is not allowed,* $(m-1)$*-veto is guaranteed to converge.*

Proof. An agent changes her preference at most once, by addition, increasing the score of one alternative she finds inferior to no other alternative. Convergence follows. □

5.54. Proposition. *If addition is not allowed,* $(m-1)$*-veto is guaranteed to converge.*

Proof. Only agents who sincerely hold a linear preference order (which they report in the first round) can make a move, incorporating flipping—a convergence proof can be obtained analogously to the proof of Proposition 5.35. □

5.55. Proposition. *If omission is not allowed,* $(m-1)$*-veto is guaranteed to converge.*

Proof. Convergence is implied by Proposition 5.35. □

5.56. Corollary. *The iterative* $(m-1)$*-veto process is guaranteed to converge if and only if omission, addition, and flipping are not simultaneously allowed.*

If only omission is allowed, convergence is immediate (holds after 0 rounds) for $(m-1)$-veto, and if addition but no flipping is allowed, convergence requires at most n rounds. Finally, if flipping is allowed (possibly together with addition or omission), the proof of Proposition 5.35 shows that we will have convergence after at most mn rounds.

Figure 5.4 provides a graphical summary of our findings. Evidently, our work raises plenty of follow-up questions: Will convergence results be affected if we consider arbitrary initial profiles (instead of only sincere ones), or random (instead of lexicographic)

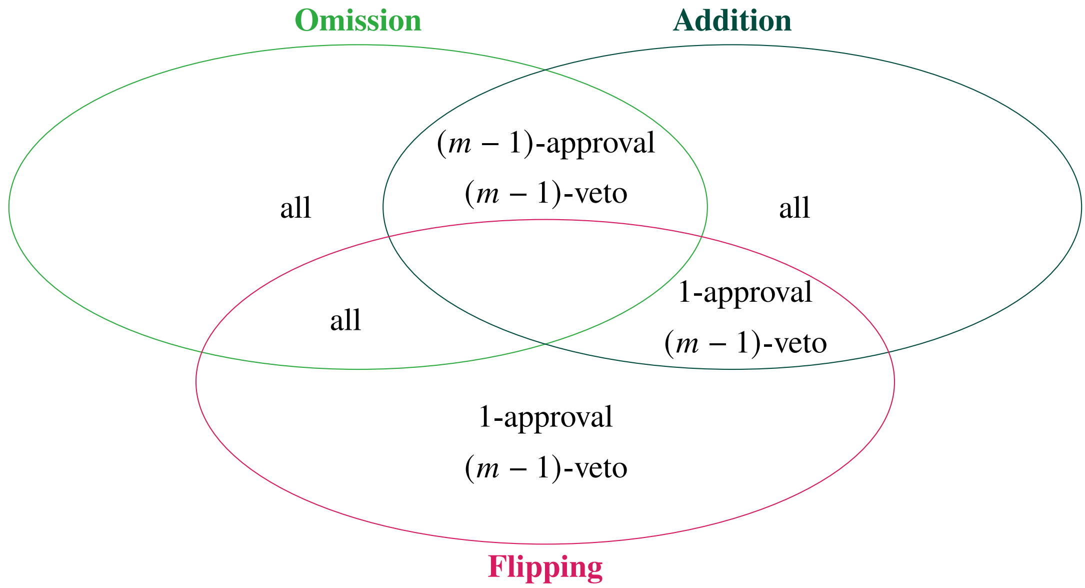


Figure 5.4: Convergence of approval and veto rules. For addition (omission) only, all rules converge. For omission and addition, $(m-1)$-approval/veto converges because flipping is needed for an agent to want to make a move twice. For flipping and omission, convergence is guaranteed for all rules besides $(m-1)$-approval because omission can do the same job as flipping, but is preferable due to being cheaper, while it cannot be repeated. For flipping and addition, "plurality-like" rules converge because the score of the winning alternative will constantly increase, reducing the set of future winners. Notably, $(m-1)$-veto converges under any condition where other rules converge too.

tie-breaking rules? Are there other sensible restrictions that we can impose to guarantee convergence? How hard (computationally) is it to calculate minimally diverging moves? What about the quality of the obtained stable states in social terms?[15] Finally, what other rules for incomplete preferences would be of interest in an iterative setting, and what if the agents can also make simultaneous updates?[16]

Questions explored in this section, in the context of voting, can also be asked within judgment aggregation (for a start regarding the complete case, see Terzopoulou and Endriss, 2018). The concepts of omission, addition, and flipping make much sense there too, and iteration might give rise to intriguing observations for various aggregation processes. However, the approval and the veto families do not have direct counterparts in judgment aggregation, so we do not foresee specific results in that territory.

[15] Related issues were addressed by Brânzei et al. (2013) in standard voting with complete preferences.
[16] Laslier (2009) studied simultaneous moves in approval voting.

5.3.3 Simulations

An important limitation of our theoretical results of Sections 5.3.1 and 5.3.2 is that they only speak about guarantees for convergence; said differently, the speak about about convergence that is achieved starting from *all* possible profiles. Analogous guarantees are reflected by results on the strategyproofness of aggregation rules, requiring that agents *never* have an incentive to misrepresent their preferences. When guarantees cannot be given, we do not yet provide any information about how *often* convergence is achieved. Furthermore, when a stable state is reached, we only know the *maximum* number of rounds needed to arrive to it—in practice, convergence can be much faster.

In this section, we complement our previous analytical results with observations that we have attained through computer simulations in Python. As we will see, convergence is essentially always achieved by the voting rules of our interest, even if we allow all possible types of manipulation (still, bearing in mind the minimal divergence assumption). This means that all stable states constitute Nash equilibria, where no agent can unilaterally benefit from changing her reported preference. For both the approval family and the veto family (particularly for $k \in \{1, m-1\}$), we explore:

- the frequency of non-manipulable sincere profiles;
- the frequency of manipulable profiles where at the end of the iterative process, when convergence is achieved, the winner is the same as that before manipulation;
- the number of rounds needed for a stable profile to be reached;
- the quality of the winner, in social terms, in sincere vs. in equilibrium profiles.

The first item above is independent of iterative voting. A number of previous works have asked the same question for voting with complete preferences, but the corresponding problem under incomplete preferences is significantly underexplored.

For profiles that consist of linear orders, Kelly (1993) conducted an experimental analysis for resolute voting rules satisfying various axiomatic properties (such as anonymity, neutrality, and a version of monotonicity) and found that the Borda rule's performance is amongst the very best ones (although the rule has been commonly criticised in terms of its susceptibility to manipulation). Kelly's experiments relied on profiles drawn uniformly at random, with at most five alternatives and three agents.

Nitzan (1985) also studied the Borda rule, as well as the plurality rule, for profiles of linear orders with $n \leqslant 90$. For both these rules, Nitzan showed that the probability of them being vulnerable to individual strategic manipulation is always greater than 0.165, 0.284, and 0.379 for three, four, and five alternatives, respectively.

In an iterative setting, Grandi et al. (2013) considered positional scoring rules for profiles of linear orders too, and conducted simulations showing that iterative manipulation under certain restrictive types of moves (that are easy to compute from the agents' perspective) in general yields a positive increase in the Condorcet efficiency of the rules (i.e., in the proportion of profiles where an alternative that wins a majority competition over all other alternatives is elected). A positive increase was also observed with respect to the group's social welfare, assuming that an agent i's cardinal utility given an alternative a is the Borda score that a gets in i's sincere preference. Grandi et al. generated profiles using the *impartial culture* (IC) assumption (implying a uniformly random distribution) and versions of the urn model. Similar observations were obtained by Reijngoud and Endriss (2012).

Meir et al. (2014) proposed the *local-dominance theory* of voting equilibria, suggesting a type of restrictions for the agents' manipulation moves that constitute a subset of their best responses. Using $n \in \{10, 20, 50\}$ and $m \in \{3, \ldots, 8\}$, they showed that the iterative plurality process quickly converges to an equilibrium starting from sincere profiles, and that the quality of the winner (again based on Borda utilities) generally improves compared to sincere voting for the IC assumption, but not for the urn model.

The IC assumption, which is employed by ample papers on voting simulations, is investigated in depth by Nurmi (1992). Broadly speaking, this assumption is useful when one thinks of the space of all possible profiles and is interested in finding the probability that a certain profile satisfies a given property—by sampling sufficiently many profiles, the desired probability is approximated. Moreover, the IC assumption is known to provide some "worst-case" guarantees for the observations obtained through simulations; for instance, it maximises the probability of cyclic collective preferences to occur under the pairwise majority rule (Tsetlin et al., 2003). Yet, people's votes in the real world cannot be expected to follow a uniform distribution.

In order to capture a certain kind of correlation between the preferences in a group, the literature has considered variations of the Polya-Eggenberger *urn model*—the one employed by Meir et al. (2014) works as follows: For a given m number of alternatives and a parameter ℓ, we start with an urn containing all possible preferences that an agent may hold (for linear orders, there are $m!$ such preferences in total—for partial orders, there are many more). We select ℓ random preferences from the urn, which will be the "leading" preferences. We make enough copies of each one of those preferences and place all of them back in the urn, so that in the end each leading preference constitutes $\frac{1}{\ell+1}$ the urn. Each agent then chooses a random preference from the urn.

For every combination of $n \in \{10, 20, 50\}$ and $m \in \{3, 4, 5\}$, we generate 200 profiles under the IC assumption and the aforementioned urn model with $\ell = 2$.[17]

[17]However, five alternatives seem to be one too many for our program to terminate under the 1-veto rule, so we skip that case (this does not prohibit us from comparing our experimental observations with our analytical results, since cycles were already observed under the 1-veto rule for four alternatives, in the proof of Proposition 5.45). Interestingly, this problem does not arise under the 1-approval rule, nor under the $(m-1)$-approval rule, which brings to the surface further, computational differences between the approval and the veto families. The difficulty linked to the iterative 1-veto process stems from the fact that the manipulation moves of the agents under that rule will often be very costly, which will make it harder, and slower, for our program to compute them. This can be seen if we consider an agent i that reports a preference $e \rhd_i d \rhd_i c \rhd_i b \rhd_i a$ in round t, and suppose that all alternatives have the same total score in that round (so a wins, because of the lexicographic tie-breaking rule). Then, agent i must reduce the score of three alternatives, b, c, and d, in order to make her favourite alternative win, which would induce cost 8 by omission. Under the plurality rule on the other hand, such situations do not emerge.

We also explicitly generate profiles of linear orders, since in our theoretical results of Section 5.3 we commonly found that cycles were formed when iteration started from sincere profiles of complete preferences—exploring such profiles through extensive simulations will give us a better idea about the role that completeness plays for convergence. Agents with complete preferences will straightforwardly have more possible incentives for manipulation, caused by the larger number of alternatives they rank.

Non-manipulable Sincere Profiles. A profile is *manipulable* under an aggregation rule if some agent has an incentive to change her preference and steer the outcome towards a winner that she prefers. Figure 5.5 depicts our experimental results concerning the (quite high) frequency of sincere profiles that are *not* manipulable, for $m = 4$.

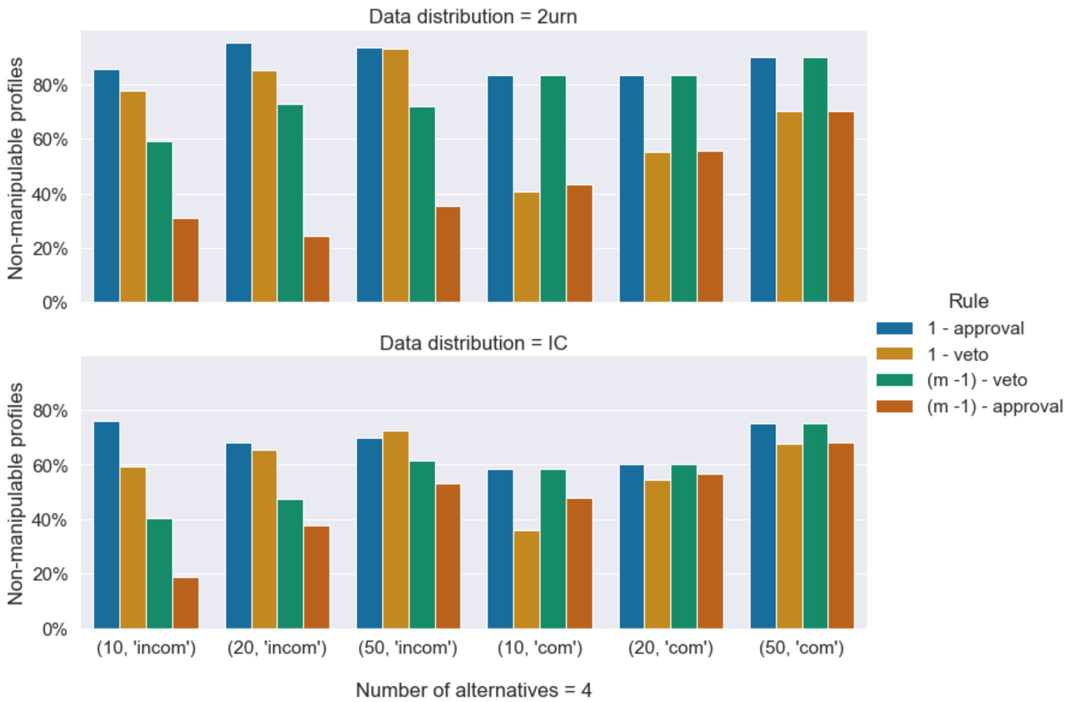


Figure 5.5: Frequency of non-manipulable sincere profiles, incomplete (‘incom’) and complete (‘com’) ones, for $n \in \{10, 20, 50\}$.

Note that the difference with respect to the frequency of manipulation between 1-approval and $(m-1)$-veto is insignificant under complete profiles, as the two rules will initially coincide and differ only in an apparently negligible manner during manipulation. We confirm three natural hypotheses: that manipulable profiles are rarer under the 2urn distribution than under the IC one,[18] as well as that complete sincere profiles and larger numbers of alternatives are detrimental to manipulability.

[18]Surprisingly, 2urn-drawn profiles are more manipulable than IC-drawn ones for $(m-1)$-approval. This contrast will further stand out later, when we talk about convergence speed—we will address it then.

Starting with incomplete profiles, we see an ordering with respect to the frequency of manipulability under the different rules for both the 2urn and the IC models—from less frequently manipulable to more frequently manipulable, the rules are: 1-approval, 1-veto, $(m-1)$-veto, and $(m-1)$-approval. Moreover, the difference between 1-approval and $(m-1)$-approval (and similarly between 1-veto and $(m-1)$-veto) is larger under incomplete initial profiles than under complete ones—possibly because there is a higher chance for ties when $k = m - 1$ under incomplete profiles.

Interestingly, we also observe that the number of non-manipulable profiles increases with the size of the group, across rules, distributions, and numbers of alternatives: Since the rules we investigate assign the same score to many alternatives in a given preference (often to all of them), they can be expected to produce many ties in small groups, implying that several agents may be able to make their favourite alternatives win.

Winner Invariance. Figure 5.6 illustrates our results regarding winner invariance for $m = 4$, but our observations below are not affected by the number of alternatives.

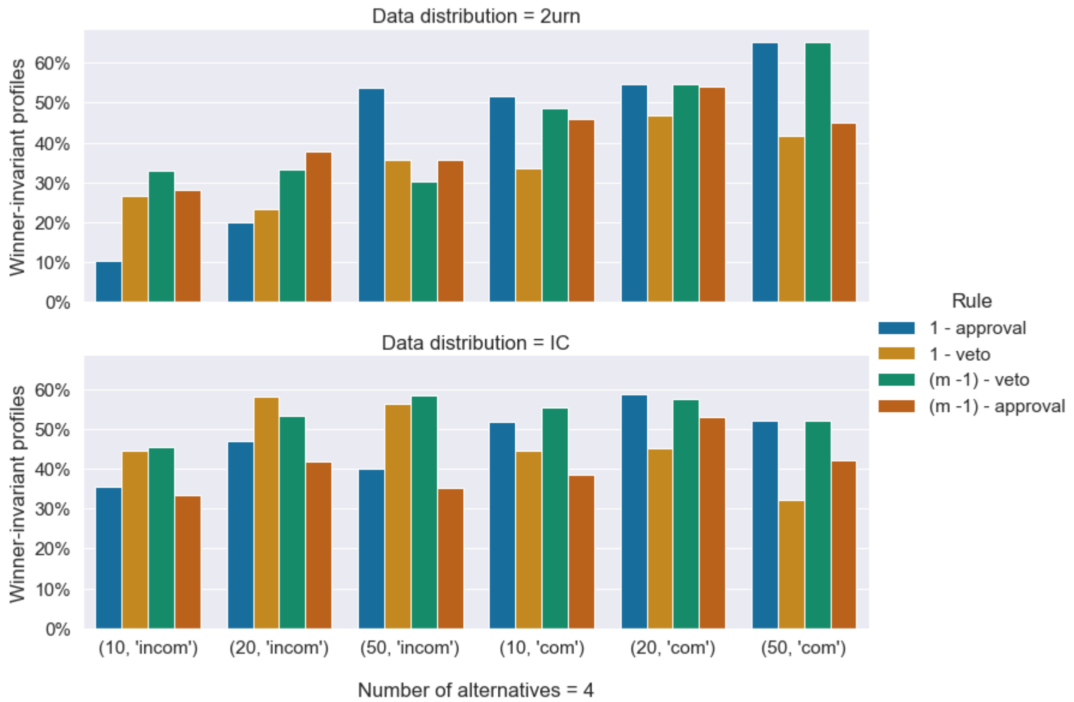


Figure 5.6: Frequency of winner-invariant profiles, incomplete ('incom') and complete ('com') ones, amongst all those that are manipulable, for $n \in \{10, 20, 50\}$.

When manipulation cannot be avoided in a given profile, it would be comforting to know that allowing the agents to manipulate in rounds has no final impact on the winner. We explore how often such a scenario arises— where the winner in equilibrium coincides with the sincere winner—bringing out positive news in various cases.

Under all rules and distributions, winner invariance is in general achieved more frequently as the group grows larger and for complete sincere profiles: Since we saw that large groups encourage non-manipulability while complete profiles damage it, we cannot deduce any direct connection between winner invariance and manipulability.

For complete initial profiles, around 50% of the iteration processes end up not changing the winner, and plurality-like rules (i.e., 1-approval and $(m-1)$-veto) perform better than antiplurality-like ones (i.e., 1-veto and $(m-1)$-approval). For incomplete initial profiles and smaller groups of agents, we detect that profiles drawn from the IC model are more frequently winner-invariant than those drawn from the 2urn model.

Speed of Convergence. Probably the most important conclusion of our simulations is that convergence is *always* achieved in practice, even if sometimes it takes a while. In the plots below (all concerning profiles with four alternatives and ten agents), we only depict up to fifty manipulation rounds. The number of alternatives does not significantly affect our observations, while larger groups—reasonably—delay convergence.[19]

Figures 5.7 and 5.8 highlight the radical, negative effect of completeness on convergence speed; a similar—albeit weaker—effect of the IC assumption follows from Figures 5.9 and 5.10 (with an exception regarding $(m-1)$-approval, which we discuss).

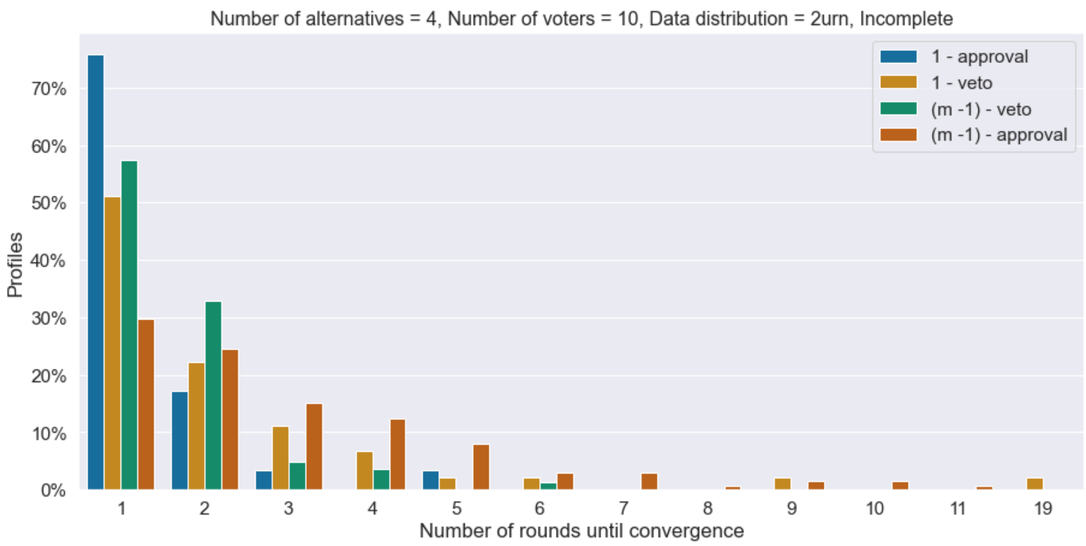


Figure 5.7: Speed of convergence (for incomplete, manipulable profiles, 2urn-drawn).

Plurality-like rules (i.e., 1-approval and $(m-1)$-veto) generally reach convergence faster than the rest—frequently, they converge after only a few rounds. Indeed, antiplurality-like rules are more sensitive to manipulation by definition: they enable the agents to manipulate by decreasing the score of the winning alternative, which means that some

[19]So in large groups, it is less common to find profiles that are manipulable, but those that are manipulable take longer to converge.

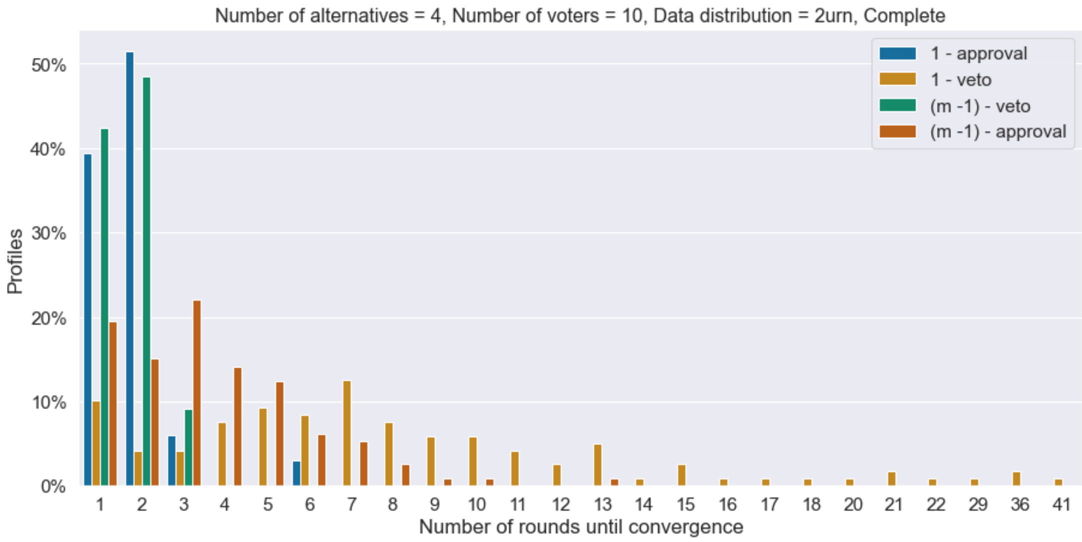


Figure 5.8: Speed of convergence (for complete, manipulable profiles, 2urn-drawn).

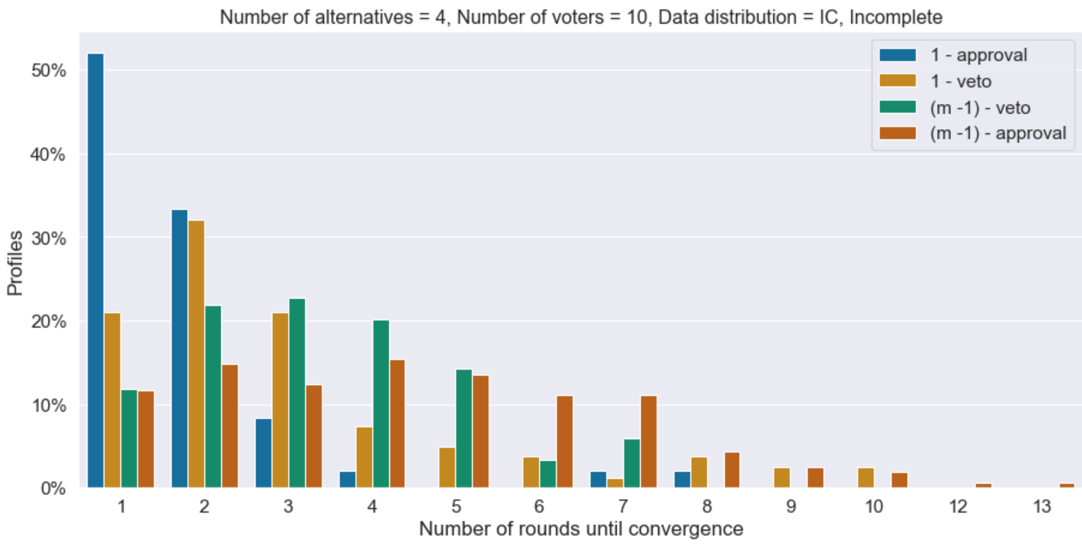


Figure 5.9: Speed of convergence (for incomplete, manipulable profiles, IC-drawn).

alternatives that have no chance of winning in the initial profile may become potential winners in future rounds (this phenomenon never emerges under the plurality-like rules, where the agents always increase the scores of alternatives by manipulating).

In the cases where the initial, sincere profiles of the agents are complete, 1-veto clearly is the slowest rule. At the same time, $(m-1)$-approval often performs quite well under the same circumstances—in particular, for complete and IC-drawn profiles, $(m-1)$-approval is amongst the fastest rules. Here is a possible interpretation for this: Agents can manipulate $(m-1)$-approval either by increasing the score of an attractive alternative, or by decreasing the score of an appalling one—they will start with the

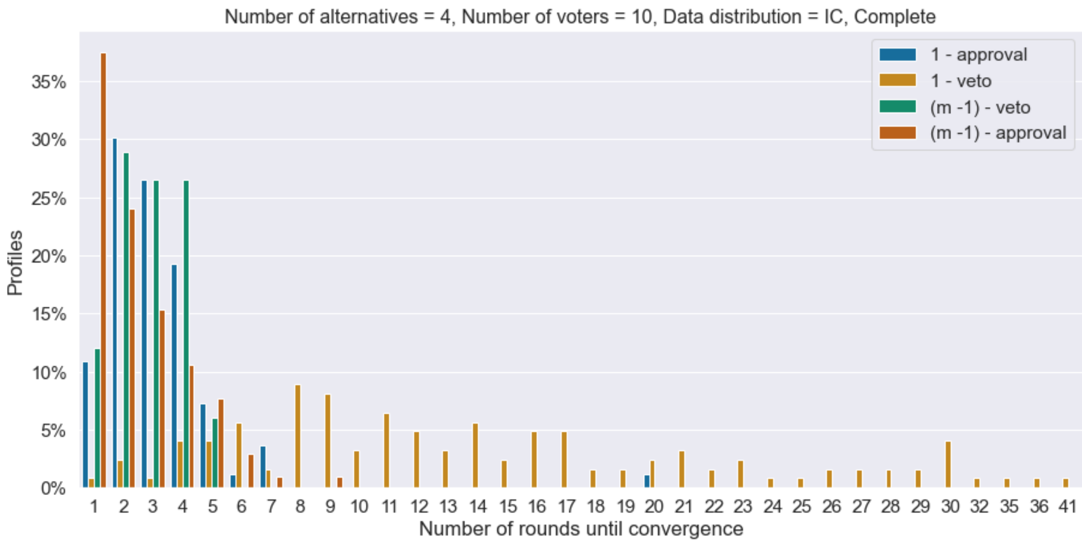


Figure 5.10: Speed of convergence (for complete, manipulable profiles, IC-drawn).

latter type of move in the initial, sincere profile. But only one alternative can have score 0 under the $(m-1)$-approval rule for complete profiles. So, a decrease in score of an alternative a will imply the increase in score of another alternative b ranked below a by the manipulator, who will only decrease the score of a if she is not in danger of making b win. In IC profiles, ties between the alternatives are more common than in 2urn profiles (because in the IC model, loosely speaking, alternatives are uniformly distributed across positions in the agents' preferences). Therefore, the probability that b will win is larger under the IC, disincentivising agents from manipulation.

Quality of Iteration. Is iteration profitable for the agents, in terms of social welfare? To answer this question we first have to define social welfare, which is not straightforward for incomplete preferences. For complete preferences, social welfare is routinely calculated using Borda scores. The idea is that the Borda score that an alternative gets from an agent's sincere preference encodes how much the agent is *relieved* by seeing that alternative win, or equivalently how much she does not *regret* seeing it win.

As we know, there is no unique way to extend Borda scores for incomplete preferences. We could examine versions of *pessimistic* and *optimistic* social welfare, corresponding to the relevant Borda generalisations, and implying that the agents are driven only by regret or only by relief, respectively. To keep things simple, here we will only exhibit results with respect to the *averaged* version of the Borda scores, meaning that our agents' utility is calculated by subtracting the regret they feel from their relief. By summing up the utilities of all agents, we obtain the social welfare of the group. The *quality of iteration* then is the improvement in social welfare from the winner in the sincere profile to the winner in the equilibrium profile.

There are several conclusions that we draw from our experiments regarding quality of iteration. Firstly, a substantial percentage of all manipulable profiles (amongst those that change the winning alternative only) do not have any impact on the agents' social welfare. This means that the group as a whole is often not affected by the consequences of iterative manipulation. When an effect on the quality of the winner arises, it takes both positive and negative values; these values are somewhat symmetric, but tend to weigh more on the positive side. Figure 5.11 (for IC-drawn, complete profiles with four alternatives and ten voters) is representative of our results.

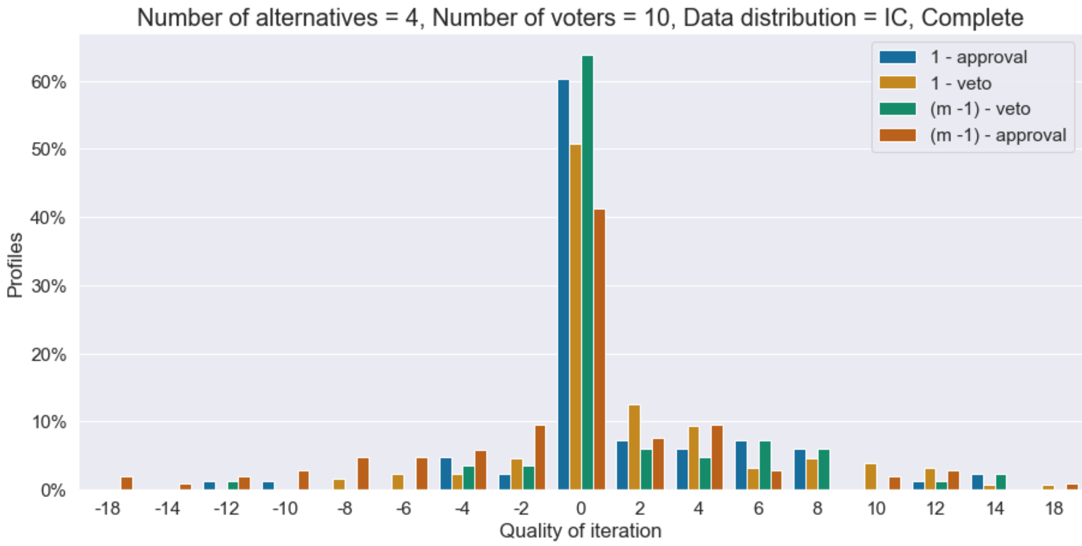


Figure 5.11: Improvement in social welfare from the sincere to the equilibrium winner (for manipulable, complete profiles, IC-drawn).

We have also observed that incomplete sincere profiles imply a wider range of values for the quality of iteration; the same holds for larger groups of agents and larger sets of alternatives. We do not include any relevant plot to demonstrate these effects here because the extreme length of their corresponding ranges makes the plots difficult to read. We have not detected any particular pattern with respect to the different rules.

To summarise, we can now confidently say that iterative manipulation processes with incomplete preferences are not so alarming as they may have seemed initially. First, very often they will not even start. Second, even if they start, they will eventually stop. Third, when they stop, they will probably do not have any impact on the winning alternative. Fourth, even if they do have an impact on the winning alternative, this will likely be a positive one for the group. All these observations agree with the ones made in the special case where the agents are only allowed to report complete preferences (recall, for instance, the work of Meir et al., 2014).

Chapter 6

Discovering a Ground Truth

Up to now, we have been studying collective decisions concerning issues that are subjective in nature, including preferences over hotels and opinions over the menu of college cafeterias. However, collective decision making often takes place for the purposes of discovering the objectively true state of a given situation. For example, doctors come together to identify the exact illness of a patient, and recruitment committees are asked to express judgments on the suitabilty of candidates for a job. It is well-known that the accuracy of a group exceeds the accuracy of a single agent, provided that all members of the group independently make better than random judgments (de Condorcet, 1785). But the (in)completeness of individual opinions constitutes a vital assumption in such contexts as well—this is what we explore in this chapter.

> Consider agents with opinions about objectively true or false issues, and with accuracy that correlates with their opinions' degree of incompleteness. What is the (probabilistically) optimal way to obtain a collective decision that discovers the correct answers on the issues at stake?

More specifically, suppose that a group of agents need to collectively determine the answer to a binary question that directly depends on the evaluation of several independent criteria. A correct *yes*/*no* answer—both on the different criteria and on the complex question—exists, but the agents are *a priori* unaware of it. Still, the agents can reflect on the possible answers and obtain a judgment which has a certain probability of being correct. But, most importantly, different agents may assess different parts of the question under consideration. We assume that, under time restrictions and cognitive constraints, the more criteria a given agent tries to assess, the less accurate her judgments are likely to be. This decrease in accuracy might be due to *time pressure* (Payne et al., 1988; Edland and Svenson, 1993; Ariely and Zakay, 2001), *multitasking attempts* (Adler and Benbunan-Fich, 2012), or *speeded reasoning* (Wilhelm and Schulze, 2002). How can the agents then, as a group, maximise the probability of discovering the correct answer to the complex question they are facing? Consider the following example, which draws inspiration from Example 3.1, at the beginning of Chapter 3, page 20.

6.1. Example. An academic hiring committee needs to decide whether Dan should get the advertised research job. In order to do so, the committee members (Ann, Bea, and Cal) have to review two of Dan's papers (p_1 and p_2)—Dan will be hired if and only if both these papers are marked as "excellent" (for the sake of the example, we assume that the excellency of an academic work as well as the suitability of a candidate for a job are objective issues). Due to an urgent deadline the committee is given only one day to judge the quality of Dan's papers. After the day passes, Ann has spent all her time on one of the two papers, while Bea and Cal have looked at both, and they express the following opinions, related to whether the relevant paper is excellent:

	p_1	p_2
Ann	☑	☐
Bea	☒	☑
Cal	☒	☑

Assuming that Ann has a higher probability to be correct than Bea about p_1, but also taking into account that Cal agrees with Bea, what is the best way to aggregate the given judgments if the committee wants to be as accurate as possible on Dan's evaluation? $\triangle$

Our analysis is situated within the framework of judgment aggregation. Along the lines of Example 6.1, we will study a special case where the propositions can be separated into the *premises* (e.g., excellency of Dan's papers) and the *conclusion* (e.g., Dan's hiring), where the conclusion is satisfied if and only if all premises are. Given two independent premises φ and ψ and a group of agents, each of which answers specific questions regarding the premises, we consider two cases of practical interest:

- *Free assignment:* Each agent chooses with some probability whether to report an opinion only on the first premise, only on the second premise, or on both.
- *Fixed assignment:* Each agent is asked (and required) to report an opinion only on the first premise, only on the second premise, or on both.

We wish to get an aggregate judgment on the conclusion that has—in expectation—high chances to reflect the truth. Assuming that the agents are sincere about the judgments they obtain after contemplating their appointed premises, we will find in Section 6.1 that the optimal aggregation rule is always a weight rule (as defined in Section 3.1).

But a further problem arises, namely that the agents might behave strategically, trying to manipulate the collective outcome to satisfy their own preferences. In Section 6.2, we will examine the three most natural cases for the preferences of an agent in our context, i.e., preferences that prioritise outcomes that are close to (i) the truth, (ii) the agent's reported judgment, or (iii) the agent's sincere judgment. In addition, we will study how an agent's incentives to be insincere relate to the information the agent holds about the judgments reported by her peers.

Returning to our assumption that the agents will be sincere, we also ask (from a mechanism-design point of view) in Section 6.3: Which fixed assignment is the most efficient one, meaning that it achieves the highest probability of producing a correct collective judgment? Our answer here depends heavily on the number of agents in the group as well as on exactly how accurate the agents are individually.

Prior work on judgment aggregation aiming at the tracking of the truth, which can be traced all the way back to the famous Condorcet Jury Theorem (de Condorcet, 1785), has primarily focused on scenarios with two independent premises and one conclusion, like the one investigated in this section. But such work has solely been concerned with the case of *complete* judgments, that is, the special case where all agents report opinions on all propositions under consideration. Under this assumption, Bovens and Rabinowicz (2006) and de Clippel and Eliaz (2015) have compared two famous aggregation rules (for uniform and varying individual accuracies, respectively): the *premise-based* rule (according to which the collective judgment on the conclusion follows from the majority's judgments on the premises) and the *conclusion-based* rule (which simply considers the opinion of the majority on the conclusion), concluding that most of the time the premise-based rule is superior. Strengthening this result, Hartmann et al. (2010) and Hartmann and Sprenger (2012) have shown that the premise-based rule is optimal across wider classes of aggregation rules too. Generalising the model further, Bozbay et al. (2014) have studied scenarios with any number of premises and agents with incentives to manipulate the collective outcome, and design rules that are optimal truth-trackers, but again assuming complete reported judgments. Also focusing on strategic agents, Ahn and Oliveros (2014) have wondered "should two issues be decided jointly by a single committee or separately by different committees?" This question differs essentially from the one addressed in our work, since our model accounts for the lower accuracy of the agents who judge a greater number of issues.

6.1 The Model and the Optimal Rule

Let φ and ψ be two logically independent premises and $c = (\varphi \wedge \psi)$ the corresponding conclusion, and assume that all three propositions are associated with a correct answer (*yes*/*no*), where a positive answer on the conclusion is equivalent to a positive answer on both premises. Each agent i in a group N holds a *sincere* judgment $J_i^* \subseteq \{\varphi, \neg\varphi, \psi, \neg\psi\}$ that contains at most one formula from each pair of a premise and its negation. Clearly, agent i cannot judge the conclusion without having judged both premises, but her judgment on the conclusion would follow directly from her judgment on the two premises in case she had one. In this setting, we denote by $\mathcal{J}$ the set of all admissible individual judgments. We say that two judgments J, J' *agree* on their evaluation of a proposition if they both contain either the non-negated or the negated version of it.

Here, we are only interested in aggregation rules F that yield consistent collective judgment sets that are complete (i.e., include an evaluation on the conclusion too, besides the two premises). We write $J^\blacktriangle \subseteq \{\varphi, \neg\varphi, \psi, \neg\psi, c, \neg c\}$ for the judgment that captures the correct evaluation on all three propositions.

We define N_1^φ (N_2^φ) to be the sets of agents who report a judgment on one (two) premise(s) and say *yes* on $\widetilde{\varphi}$. We define $n_1^\varphi = |N_1^\varphi|$ and $n_2^\varphi = |N_2^\varphi|$ to be the corresponding cardinalities of these sets. Analogously, we define the sets of agents who say *no* on $\widetilde{\varphi}$ by replacing the proposition φ with $\neg\varphi$ in the aforementioned notation.

We denote by $\mathtt{p}$ the probability that agent i's sincere judgment J_i^* is correct on a premise when i judges both premises and by $\mathtt{q}$ the relevant probability when i only judges a single premise (assuming that the probability of each agent i's judgment being correct on a premise φ is independent (i) of whether φ is true or false and (ii) of what i's judgment on premise ψ is). We assume that the probabilities $\mathtt{p}$ and $\mathtt{q}$ are the same for all agents, but the agents make their judgments independently of each other. We shall moreover suppose that all agents' judgments are more accurate than a random guess, but not perfect, and that agents judging a single premise are strictly more accurate than those judging both premises, i.e., that $1/2 < \mathtt{p} < \mathtt{q} < 1$. Then, $P(\boldsymbol{J}^*)$ denotes the probability of the sincere profile $\boldsymbol{J}^*$ to arise and $P(\boldsymbol{J}^*_{-i} \mid J_i^*)$ the probability that the judgments of all agents besides i form the sincere (partial) profile $\boldsymbol{J}^*_{-i}$, given that i has the sincere judgment J_i^*. Formally, for a fixed assignment it holds that:

$$P(\boldsymbol{J}^*) = P(\varphi \text{ true}) \cdot P(\boldsymbol{J}^* \mid \varphi \text{ true}) + P(\varphi \text{ false}) \cdot P(\boldsymbol{J}^* \mid \varphi \text{ false})$$

where $P(\boldsymbol{J}^* \mid \varphi \text{ true}) = \mathtt{q}^{n_1^\varphi} \mathtt{p}^{n_2^\varphi} (1-\mathtt{q})^{n_1^{\neg\varphi}} (1-\mathtt{p})^{n_2^{\neg\varphi}}$, and similarly for $P(\boldsymbol{J}^* \mid \varphi \text{ false})$. The *accuracy* $P(F)$ of a resolute aggregation rule F is defined as follows:

$$P(F) \quad = \sum_{\substack{\boldsymbol{J}^* \in \mathcal{J}^n \text{ s.t.} \\ F(\boldsymbol{J}^*) \text{ and } J^\blacktriangle \text{ agree on } c}} P(\boldsymbol{J}^*)$$

We next define the (irresolute) aggregation rule F_{opt}^{irr}: it functions as a weighted-majority rule on each premise separately, assigning to the agents weights according to the size of their reported judgments, and subsequently picks that evaluation of the conclusion that is consistent with the collective evaluation of the premises—this is exactly a weight rule, as defined in Section 3.1. More specifically, we define the following weights:

$$w_i = \begin{cases} \log \frac{\mathtt{q}}{1-\mathtt{q}} & \text{if } |J_i| = 1 \\ \log \frac{\mathtt{p}}{1-\mathtt{p}} & \text{if } |J_i| = 2 \\ 0 & \text{if } |J_i| = 0 \end{cases}$$

Observe that the base of the logarithm in the definition of w_i above is irrelevant. Then, for all profiles $\boldsymbol{J}$ we have the following:

$$F_{opt}^{irr}(\boldsymbol{J}) = \underset{J \in \mathcal{J}(\Phi)^\bullet}{\operatorname{argmax}} \sum_{i \in N} w_i \cdot |J \cap J_i|$$

F_{opt} is a resolute version of F_{opt}^{irr} that, if the obtained collective judgments are more than one, randomly chooses one of them for the collective outcome.

For a resolute aggregation rule F, the probability $P(F)$ depends on the probabilities $P(F$ correct on $\varphi)$ and $P(F$ correct on $\psi)$, which, for simplicity, we call P_φ and P_ψ, respectively. For the remainder of this chapter we will further assume that the prior probabilities of the two premises being true or false are equal (and independent of each other). That is, $P(\varphi \text{ true}) = P(\psi \text{ true}) = 1/2$. Then:

$$\begin{aligned} P(F) &= \frac{1}{4}\Big[\big(P_\varphi P_\psi\big) + \big(P_\varphi P_\psi + (1-P_\varphi)(1-P_\psi) + P_\varphi(1-P_\psi)\big) + \big(P_\varphi P_\psi + \neg\Phi \\ &\quad (1-P_\varphi)(1-P_\psi) + (1-P_\varphi)P_\psi\big) + \big(P_\varphi P_\psi + P_\varphi(1-P_\psi) + (1-P_\varphi)P_\psi\big)\Big]\neg\Phi \\ &= \frac{1}{2} + \frac{P_\varphi P_\psi}{2} \end{aligned} \tag{6.1}$$

Given a (fixed or free) assignment, let us denote by $\boldsymbol{J}_F^\varphi$ ($\boldsymbol{J}_F^{\neg\varphi}$) and $\boldsymbol{J}_F^\psi$ ($\boldsymbol{J}_F^{\neg\psi}$) the sets of all possible profiles of reported judgments that lead to a *yes* (*no*) collective answer on φ and ψ under the rule F, respectively. Then:

$$P_\varphi = \frac{1}{2}\sum_{\boldsymbol{J}^* \in \boldsymbol{J}_F^\varphi} P(\boldsymbol{J}^* \mid \varphi \text{ true}) + \frac{1}{2}\sum_{\boldsymbol{J}^* \in \boldsymbol{J}_F^{\neg\varphi}} P(\boldsymbol{J}^* \mid \varphi \text{ false}) \tag{6.2}$$

Now, for a fixed assignment and a profile $\boldsymbol{J}^*$, we have the following:

$$\begin{array}{rcll} P(\boldsymbol{J}^* \mid \varphi \text{ true}) & > & P(\boldsymbol{J}^* \mid \varphi \text{ false}) & \Leftrightarrow \\ \mathsf{q}^{n_1^\varphi}\mathsf{p}^{n_2^\varphi}(1-\mathsf{q})^{n_1^{\neg\varphi}}(1-\mathsf{p})^{n_2^{\neg\varphi}} & > & (1-\mathsf{q})^{n_1^\varphi}(1-\mathsf{p})^{n_2^\varphi}\mathsf{q}^{n_1^{\neg\varphi}}\mathsf{p}^{n_2^{\neg\varphi}} & \Leftrightarrow \\ n_1^\varphi \log\dfrac{\mathsf{q}}{1-q} + n_2^\varphi \log\dfrac{\mathsf{p}}{1-p} & > & n_1^{\neg\varphi}\log\dfrac{\mathsf{q}}{1-\mathsf{p}} + n_2^{\neg\varphi}\log\dfrac{\mathsf{p}}{1-\mathsf{p}} & \end{array} \tag{6.3}$$

Analogously, we consider a free assignment where agent i makes a sincere judgment on $\widetilde{\varphi}$ with probability $p_i^{\widetilde{\varphi}}$, on $\widetilde{\psi}$ with probability $p_i^{\widetilde{\psi}}$, and on both premises with probability $p_i^{\widetilde{\varphi},\widetilde{\psi}}$. Given a sincere profile $\boldsymbol{J}^*$, the following holds:

$$P(\boldsymbol{J}^* \mid \varphi \text{ true}) = \mathsf{q}^{n_1^\varphi}\mathsf{p}^{n_2^\varphi}(1-\mathsf{q})^{n_1^{\neg\varphi}}(1-\mathsf{p})^{n_2^{\neg\varphi}} \prod_{i \in N_1^\varphi \cup N_1^{\neg\varphi}} p_i^{\widetilde{\varphi}} \prod_{j \in N_2^\varphi \cup N_2^{\neg\varphi}} p_j^{\widetilde{\varphi},\widetilde{\psi}}$$

Defining $P(\boldsymbol{J}^* \mid \varphi \text{ false})$ similarly, we have the following as before:

$$\begin{array}{rcll} P(\boldsymbol{J}^* \mid \varphi \text{ true}) & > & P(\boldsymbol{J}^* \mid \varphi \text{ false}) & \Leftrightarrow \\ n_1^\varphi \log\dfrac{\mathsf{q}}{1-\mathsf{q}} + n_2^\varphi \log\dfrac{\mathsf{p}}{1-\mathsf{p}} & > & n_1^{\neg\varphi}\log\dfrac{\mathsf{q}}{1-\mathsf{q}} + n_2^{\neg\varphi}\log\dfrac{\mathsf{p}}{1-\mathsf{p}} & \end{array} \tag{6.4}$$

Theorem 6.2 states the main results of this section. Its proof technique is a standard method in research on *maximum likelihood estimators* (Dawid and Skene, 1979).

6.2. Theorem. *For any fixed (or free) assignment and sincere judgments, it holds that* $F_{opt} \in \operatorname{argmax}_F P(F)$. *For every other aggregation rule* $F' \in \operatorname{argmax}_F P(F)$, F' *only differs from* F_{opt} *on the tie-breaking part.*

Proof. For a fixed assignment, it follows from Equations (6.2) and (6.3) that P_φ (and P_ψ) will be maximal if and only if F assigns to the agents weights as in F_{opt}. Equation (6.1) implies that $\max_F P(F) \leqslant \frac{1}{2} + \frac{\max_F P_\varphi \max_F P_\psi}{2}$, so $P(F)$ is maximal if and only if F_{opt} (or a rule that only differs from F_{opt} on the tie-breaking part) is used. The proof is analogous for a free assignment, using Equation (6.4). □

Let us illustrate how the optimal aggregation rule works with some simple examples, building on the story of Example 6.1 and considering specific accuracies p and q.

6.3. Example. Recall Example 6.1, and suppose that $\mathtt{q} = 0.8$ and $\mathtt{p} = 0.6$. A *yes* on proposition p_1 will get weight $\log \frac{0.8}{0.2} \simeq 0.6$, from Ann, while a *no* on the same proposition will be assigned weight $2 \log \frac{0.6}{0.4} \simeq 0.34$, from Bea and Cal. The former number is larger than the latter, so the collective decision will be in favour of p_1 (and in favour of p_2, since there is unanimous agreement for it). Dan can get the job. △

6.4. Example. Again in Example 6.1, suppose that $\mathtt{q} = 0.65$ and $\mathtt{p} = 0.6$. A *yes* on proposition p_1 will get weight $\log \frac{0.65}{0.35} \simeq 0.27$, from Ann, while a *no* on the same proposition will get weight $2 \log \frac{0.6}{0.4} \simeq 0.34$, from Bea and Cal. The former number is smaller than the latter, so the collective decision will be against p_1 (but in favour of p_2, since there is unanimous agreement for it). Dan will not get the job. △

6.2 Strategic Manipulation

In this section we study the incentives of the agents to report insincere judgments when the most accurate rule F_{opt} is used. We examine in detail both fixed and free assignments. An agent's incentives to be insincere will of course depend on the type of her preferences. We analyse three natural cases regarding these preferences, assuming that all agents have the same preference type: First, the agents may want the group to reach a correct judgment—these preferences are called *truth-oriented*. Second, the agents may want to report an opinion that is close to the collective judgment, whatever that judgment is—these preferences are called *reputation-oriented*. Third, the agents may want the group's judgment to agree with their own sincere judgment—these preferences are called *self-oriented*.[1] Recall that the agents' opinions in this chapter are only taken to conern the premises, and not the conclusion. So, our definitions of the agents' preferences will not consider the conclusion either.

[1]For instance, doctors making judgments about their patients may simply care about the correctness of their collective judgment, participants of an experiment that are paid proportionally to their agreement with the group can be assumed to aim at being seen to agree with their peers, and people who like having their opinions confirmed might manipulate the group to agree with their own privately held judgment.

To make things formal, we employ tools from Bayesian game theory. We wish to understand when sincerity by all agents is an equilibrium:

> Given that agent i holds the sincere judgment J_i^*, and given that the rest of the agents are going to be sincere no matter what judgments they have, is sincerity (i.e., reporting J_i^*) a best response of agent i?

We examine the *interim* and the *ex-post* case. In both cases agent i already knows her own sincere judgment, but in the former case she is ignorant about the judgments of the rest of the group (only knowing that they have to be probabilistically compatible with her own judgment), while in the latter case she is in addition fully informed about them (this can happen, for example, after some communication action has taken place).

Call $T \in \{$"*truth*", "*reputation*", "*self*"$\}$ the type of the agents' preferences. Let us denote by $U_i^T\big((J_i, \boldsymbol{J}_{-i}^*), \boldsymbol{J}^*\big)$ the utility that agent i gets by reporting judgment J_i, when the sincere profile of the group is $\boldsymbol{J}^*$ and all other agents $j \neq i$ report their sincere judgments J_j^*.[2] $EU_i^T\big((J_i, \boldsymbol{J}_{-i}^*), J_i^*\big)$ stands for the expected utility that agent i gets by reporting judgment J_i, when her sincere judgment is J_i^* and all other agents j report their sincere judgments for any possible such judgments. More precisely:

$$
\begin{aligned}
U_i^{truth}\big((J_i, \boldsymbol{J}_{-i}^*), \boldsymbol{J}^*\big) &= |F_{opt}(J_i, \boldsymbol{J}_{-i}^*) \cap J^{\blacktriangle}| \\
U_i^{reputation}\big((J_i, \boldsymbol{J}_{-i}^*), \boldsymbol{J}^*\big) &= |F_{opt}(J_i, \boldsymbol{J}_{-i}^*) \cap J_i| \\
U_i^{self}\big((J_i, \boldsymbol{J}_{-i}^*), \boldsymbol{J}^*\big) &= |F_{opt}(J_i, \boldsymbol{J}_{-i}^*) \cap J_i^*|
\end{aligned}
$$

Also, for any $T \in \{$"*truth*", "*reputation*", "*self*"$\}$, we define the following:

$$
EU_i^T\big((J_i, \boldsymbol{J}_{-i}^*), J_i^*\big) = \sum_{\boldsymbol{J}_{-i}^*} U_i^T\big((J_i, \boldsymbol{J}_{-i}^*), \boldsymbol{J}^*\big) P(\boldsymbol{J}_{-i}^* \mid J_i^*)
$$

We proceed with formally defining *strategyproofness* in our framework, namely the situation where all agents being sincere forms an equilibrium.

6.5. Definition. *Given a preference type $T \in \{$"truth", "reputation", "self"$\}$ and a (fixed or free) assignment, we say that* ***sincerity always gives rise to an interim equilibrium*** *if and only if for all agents $i \in N$ and sincere judgments $J_i^* \in \mathcal{J}$ the following holds, where $S_i \subseteq \mathcal{J}$ is the set of all judgments that agent i is can potentially report under the given assignment:*

$$
J_i^* \in \operatorname*{argmax}_{J_i \in S_i} EU_i^T\big((J_i, \boldsymbol{J}_{-i}^*), J_i^*\big)
$$

[2]Note that, although we refer to U_i^T as the *utility* of agent i, when the definition of U_i^T for a specific type of preferences T makes use of the rule F_{opt} (which is paired with a random tie-braking rule in order to be resolute), U_i^T more precisely corresponds to the *expected utility* of agent i, since there is uncertainty regarding the collective outcome produced by F_{opt}. Nonetheless, we slightly abuse the terminology in order to be able to distinguish more clearly the situations where the expected utility of agent i is linked to agent i's uncertainty about the sincere judgments of her peers.

Similarly, *sincerity always gives rise to an ex-post equilibrium* if and only if the condition of Definition 6.5 holds, where EU_i^T is replaced by U_i^T.

Table 6.1 summarises our results, where "✓" stands for strategyproofness and "✗" designates the existence of a counterexample.

Preferences \ Assignment	Fixed *interim*	Fixed *ex-post*	Free *interim*	Free *ex-post*
truth-oriented	✓ Thm 6.8	✓ Thm 6.8	✓ Thm 6.8	✓ Thm 6.8
reputation-oriented	✓Thm 6.10	✗ Prop 6.11	✓ Thm 6.12	✗ Prop 6.11
self-oriented	✓Thm 6.14	✓ Thm 6.14	✗ Prop 6.15	✗ Prop 6.15

Table 6.1: Strategyproofness results in truth-tracking contexts.

Two fundamental lemmas are in order. First, we verify the basic intuition that when an agent holds more information about the reported judgments of the rest of the group, then her incentives to manipulate increase.

6.6. Lemma. *For any assignment and type of preferences, ex-post strategyproofness implies interim strategyproofness.*

Proof. We show the contrapositive. Suppose *interim* strategyproofness is violated. Then, there exists an agent i with a sincere judgment J_i^* such that the following holds:

$$J_i^* \notin \operatorname*{argmax}_{J_i \in A_i} EU_i^T\left((J_i, \boldsymbol{J}_{-i}^*), J_i^*\right)$$

This also means that the following holds:

$$J_i^* \notin \operatorname*{argmax}_{J_i \in A_i} \sum_{\boldsymbol{J}_{-i}^*} U_i^T\left((J_i, \boldsymbol{J}_{-i}^*), \boldsymbol{J}^*\right) P(\boldsymbol{J}_{-i}^* \mid J_i^*)$$

Hence, there must exist an insincere judgment J_i and a partial profile of judgments $\boldsymbol{J}_{-i}^*$ with $P(\boldsymbol{J}_{-i}^* \mid J_i^*) > 0$ for which the following holds:

$$U_i^T\left((J_i, \boldsymbol{J}_{-i}^*), \boldsymbol{J}^*\right) > U_i^T\left((J_i^*, \boldsymbol{J}_{-i}^*), \boldsymbol{J}^*\right)$$

But this means that we have a scenario violating *ex-post* strategyproofness. □

Second, we stress that every counterexample of *ex-post* strategyproofness under fixed assignments is a counterexample under free assignments too (since the definition of *ex-post* strategyproofness implies that the agents know each other's reported judgments).[3]

6.7. Lemma. *For any type of preferences, ex-post strategyproofness under free assignments implies ex-post strategyproofness under fixed assignments.*

[3] The other direction does not hold. Importantly, a counterexample may go through under free but not fixed assignments because the agents have the option to manipulate by abstaining on some premise they have sincerely thought about.

Truth-oriented Preferences. When all agents have truth-oriented preferences and when the rule F_{opt} is used to aggregate their reported judgments, it directly is in everyone's best interest to be sincere—given that the rest of the group is sincere as well—irrespective of whether the assignment materialised is fixed or free and whether the agents know the judgments of their peers. Intuitively, the agents can trust that the rule F_{opt} will achieve a collective judgment that is as accurate as possible.

6.8. THEOREM. *For any fixed (or free) assignment and truth-oriented preferences:*

(i) *sincerity always gives rise to an interim equilibrium*

(ii) *sincerity always gives rise to an ex-post equilibrium*

PROOF. By Lemma 6.6, we only need to prove case (ii). For an arbitrary sincere profile $\boldsymbol{J}^* = (J_i^*, \boldsymbol{J}_{-i}^*)$, we have that $U_i^{truth}\big((J_i, \boldsymbol{J}_{-i}^*), \boldsymbol{J}^*\big) = |F_{opt}(J_i, \boldsymbol{J}_{-i}^*) \cap J^\blacktriangle|$, where $J^\blacktriangle$ captures the true evaluation of the propositions. Suppose, aiming for a contradiction, that for some insinsere judgment J_i of agent i the following holds:

$$|F_{opt}(J_i, \boldsymbol{J}_{-i}^*) \cap J^\blacktriangle| > |F_{opt}(J_i^*, \boldsymbol{J}_{-i}^*) \cap J^\blacktriangle|$$

This means that $F_{opt}(J_i, \boldsymbol{J}_{-i}^*) \neq F_{opt}(J_i^*, \boldsymbol{J}_{-i}^*)$. Then, for a suitable aggregation rule $F' \neq F_{opt}$, we can write $F_{opt}(J_i, \boldsymbol{J}_{-i}^*) = F'(J_i^*, \boldsymbol{J}_{-i}^*)$ and derive the following:

$$|F(J_i^*, \boldsymbol{J}_{-i}^*) \cap J^\blacktriangle| > |F_{opt}(J_i^*, \boldsymbol{J}_{-i}^*) \cap J^\blacktriangle|$$

This is impossible, because F_{opt} has to maximise aggrement with $J^\blacktriangle$ (Theorem 6.2). Hence, it holds that $|F_{opt}(J_i, \boldsymbol{J}_{-i}^*) \cap J^\blacktriangle| \leqslant |F_{opt}(J_i^*, \boldsymbol{J}_{-i}^*) \cap J^\blacktriangle|$ for all J_i, which implies that $U_i^{truth}\big((J_i, \boldsymbol{J}_{-i}^*), \boldsymbol{J}^*\big) \leqslant U_i^{truth}\big((J_i^*, \boldsymbol{J}_{-i}^*), \boldsymbol{J}^*\big)$ for all J_i and concludes the proof. □

Example 6.9 illustrates the idea of strategyproofness under truth-oriented preferences.

6.9. EXAMPLE. Consider again the setting of Example 6.1, and suppose that Ann, Bea, and Cal provide the following judgments:

	p_1	p_2
Ann	☑	☑
Bea	☒	☑
Cal	☒	☑

Ann has a positive opinion on both p_1 and p_2. Suppose she cares about the collective decision being *correct* on p_1. What are her options, then, regarding the opinions she can report, knowing that her personal evaluation on p_1 has a higher probability of being correct than of being wrong?

Firstly, there is no reason for Ann to submit a negative judgment on p_1, because F_{opt} will then give more weight to an opinion that is less probably correct than her sincere one. She could try to be negative on p_2, but this has no chance of influencing the outcome on p_1. She can also abstain on p_2 (if she is permitted to do so, in a free assignment), in order for her judgment on p_1 to get more weight by F_{opt}. If that action changes the collective outcome on p_1, then it means that a different aggregation rule could have obtained that new outcome in the original profile (just by counting the opinions of the agents differently). Since such an aggregation rule is not the optimal one by definition, the new outcome that Ann achieves cannot be closer to the truth, meaning that Ann's preference is not satisfied. △

Reputation-oriented Preferences. When the agents care about the positive reputation they obtain by agreeing with the collective judgment of the group, their incentives to behave insincerely heavily depend on whether they already know the judgments of their peers. Of course: If an agent knows precisely what the collective judgment of the group will be, she can simply change her reported judgment to fully match that collective judgment. On the other hand, we will see that if an agent does not know exactly what the sincere judgments of her peers are, it is more attractive for her to remain sincere (as—knowing that her sincere judgment is more accurate than random—she can reasonably expect the group to agree with her).

6.10. Theorem. *For any fixed assignment and reputation-oriented preferences, sincerity always gives rise to an interim equilibrium.*

Proof. Given a fixed assignment, an agent i, and a sincere judgment J_i^*, let us call P_{dis} the probability that agent i will disagree with the group on the evaluation of premise φ. Let us assume that agent i's judgment J_i^* concerns both premises φ and ψ (the proof is analogous when J_i^* concerns only premise φ). Now, let us denote by P_g the probability that the group is collectively correct on their evaluation of φ. Recalling that $\mathrm{p} > 1/2$ is the probability that agent i is correct on her evaluation of φ, the following holds:

$$P_{dis} \leqslant \mathrm{p}(1 - P_g) + (1 - \mathrm{p})P_g = \mathrm{p} + P_g(1 - 2\mathrm{p})$$

Since all agents are more accurate than random, we also have that $P_g > 1/2$, and:

$$\mathrm{p} + P_g(1 - 2\mathrm{p}) \leqslant 1/2 \quad \text{if and only if} \quad P_g \geqslant \frac{\mathrm{p} - 1/2}{2\mathrm{p} - 1} = 1/2$$

So $P_{dis} \leqslant 1/2$, which means that it is more probable for the group's judgment to agree with J_i^* on premise φ than to disagree with it, and the same holds for premise ψ. Therefore, agent i has no better option than to report her sincere judgment on her assigned premises. Formally, $EU_i^{reputation}\big((J_i, \boldsymbol{J}_{-i}^*), J_i^*\big)$ is maximised for $J_i = J_i^*$. □

6.11. Proposition. *For reputation-oriented preferences, there exists a fixed (and thus a free, from Lemma 6.7) assignment where sincerity does not always give rise to an ex-post equilibrium.*

Proof. Consider a fixed assignment where all agents in the group are asked about both premises φ and ψ, agent i has the sincere judgment $J_i^* = \{\varphi, \psi\}$, and all other agents j have sincere judgments $J_j = \{\varphi, \neg\psi\}$. Agent i would increase her utility by reporting the insincere judgment $J_i = \{\varphi, \neg\psi\}$. □

6.12. Theorem. *For any free assignment and reputation-oriented preferences, sincerity always gives rise to an interim equilibrium.*

Proof. Consider an arbitrary free assignment and an agent i with sincere judgment J_i^*. Since the given assignment is uncertain, agent i can potentially report any judgment set she wants. Suppose that her sincere judgment J_i^* has size $|J_i^*| = k \in \{1, 2\}$. First, following the same argument as that in the proof of Theorem 6.10, we see that agent i cannot increase her expected utility by reporting a judgment $J_i \neq J_i^*$ with $|J_i| = k$. We omit the formal details, but the intuition is clear: the group has higher probability to agree with the sincere evaluation of agent i on each premise than to disagree with it.

However, we also need to show that agent i cannot increase her expected utility by reporting a judgment $J_i \neq J_i^*$ with $|J_i| \neq k$. The case where $|J_i| > k$ is straightforward: if agent i has no information about one of the premises, the best she could do is reporting a random judgment on that premise, but this would not increase her expected utility. Thus, we need to consider the case where $|J_i| < k$, and more specifically the only interesting scenario with $|J_i^*| = 2$ and $|J_i| = 1$. Say, without loss of generality, that $J_i = \{\varphi\}$. Let us call $P^{\varphi}_{ag,2}$ ($P^{\psi}_{ag,2}$) the probability that the group will agree with agent i on her evaluation of φ (ψ) given that agent i reports her sincere judgment on both premises, and let us call $P^{\varphi}_{ag,1}$ the probability that the group will agree with agent i on her evaluation of φ given that she reports her a judgment only on premise φ. Then:

$$EU_i^{reputation}\big((J_i^*, \boldsymbol{J}_{-i}^*), J_i^*\big) = P^{\varphi}_{ag,2} + P^{\psi}_{ag,2} \qquad \text{and}$$
$$EU_i^{reputation}\big((J_i, \boldsymbol{J}_{-i}^*), J_i^*\big) = P^{\varphi}_{ag,1}$$

Analogously to the proof of Theorem 6.10, it holds that $P^{\varphi}_{ag,2} \geqslant 1/2$ and $P^{\psi}_{ag,2} \geqslant 1/2$, so $P^{\varphi}_{ag,2} + P^{\psi}_{ag,2} \geqslant 1$. This means that $P^{\varphi}_{ag,1} \leqslant P^{\varphi}_{ag,2} + P^{\psi}_{ag,2}$ (because $P^{\varphi}_{ag,1} \leqslant 1$ is a probability value), so, it is the case that $EU_i^{reputation}\big((J_i, \boldsymbol{J}_{-i}^*), J_i^*\big) \leqslant EU_i^{reputation}\big((J_i^*, \boldsymbol{J}_{-i}^*), J_i^*\big)$. Hence, agent i cannot increase her expected utility by reporting J_i instead of J_i^*. □

Example 6.13 serves as a demonstration for reputation-oriented preferences.

6.13. Example. Take the profile we also saw in Example 6.9:

	p_1	p_2
Ann	☑	☑
Bea	☒	☑
Cal	☒	☑

Suppose Ann is aware of the judgments of her peers. Then, she will immediately realise that the collective decision on p_1 will be different than her sincere opinion (because in a complete profile, F_{opt} behaves as a simple majority rule). If she cares about her reputation, she can simply report an insincere judgment which would agree with the the outcome, without any risk. However, in case Ann does not exactly know the judgments of the other members in her group, such an insincere act would not be a safe one. △

Self-oriented Preferences. Suppose the agents would like the collective outcome to agree with their own sincere judgment. Now, having a fixed or a free assignment radically changes their strategic considerations: under fixed assignments the agents can never increase their utility by lying, while there are free assignments where insincere behaviour is profitable. The critical difference is that when the agents are free to submit judgments of variable size, they can increase the weight that the optimal aggregation rule F_{opt} will assign to their judgment on one of the two premises by avoiding to report a judgment on the other premise, thus having more opportunities to manipulate the outcome in favour of their private judgment.

6.14. Theorem. *For any fixed assignment and self-oriented preferences:*

(i) *sincerity always gives rise to an interim equilibrium*

(ii) *sincerity always gives rise to an ex-post equilibrium*

Proof. By Lemma 6.6, we only need to show case (ii). Given a fixed assignment, an agent can only report an insincere opinion by flipping her sincere judgment on some of the premises she is asked about. But if she did so, F_{opt} could only favour a judgment different from her own, not increasing her utility. Thus, every agent always maximises her utility by being sincere. □

6.15. Proposition. *For self-oriented preferences, there exists a free assignment such that sincerity does not always give rise to an interim equilibrium (and thus also not to an ex-post equilibrium, from Lemma 6.6).*

Proof. Consider a group of three agents and a free assignment as follows: Agent 1 reports an opinion on both premises φ, ψ with probability $1/2$, and only on premise φ or only on premise ψ with probability 1/4 and 1/4, respectively. Agent 2 reports a judgment on both premises φ, ψ with probability 1, and agent 3 reports a judgment only on premise φ with probability 1. Suppose that agent 1's truthful judgment is $J_1^* = \{\varphi, \psi\}$. For the calculation of agent 1's expected utility, note that agent 1 is correct on her evaluation of a proposition premise φ with probability $\mathtt{p}$ (and the same for premise ψ). Knowing what agent 1's sincere judgment is, we also know the probability of occurrence of every possible partial profile of judgments (J_2, J_3). For instance the probability that both agents 2 and 3 will report judgments that agree on premise φ with J_1^* is $\mathtt{p}(\mathtt{pq}) + (1-\mathtt{p})(1-\mathtt{p})(1-\mathtt{q})$.

Suppose additionally that $\mathtt{q} > \frac{\mathtt{p}^2}{\mathtt{p}^2+(1-\mathtt{p})^2}$. In such a case, if agent 1 decides to report her sincere judgment on both premises, she will always be unable to affect the collective outcome on φ according to the rule F_{opt}, and she will obtain an outcome that agrees with her sincere judgment on ψ with the following probability:

$$\mathtt{p}(\mathtt{p} + \frac{1-\mathtt{p}}{2}) + (1-\mathtt{p})(1-\mathtt{p}+\frac{\mathtt{p}}{2}) = \mathtt{p}^2 + \mathtt{p} + 1 < 1$$

However, if agent 1 reports the insincere judgment $J_1 = \{\psi\}$ instead, she will always obtain a collective outcome on ψ that is identical to her own sincere judgment, corresponding to a higher expected utility of value 1. Thus, we can conclude that $J_1^* \notin \mathrm{argmax}_{J_1 \in S_1} EU_1^{\mathrm{self}}\big((J_1, J_2^*, J_3^*), J_1^*\big)$. □

Example 6.16 instantiates strategyproofness with respect to self-oriented preferences.

6.16. Example. Consider again the profile appearing in Examples 6.9 and 6.13. If Ann has self-oriented preferences, she would like the collective decision on p_1 to agree with her sincere judgment, which would not be the case if she remains sincere. Also, Ann could not direct the outcome towards her sincere judgment by reporting an opposite opinion on p_1, by construction of the rule F_{opt} (that works in a majority-style). However, Ann can insincerely attract more weight on her judgment by pretending she holds a smaller number of concrete opinions (and thus F_{opt} would deem her more accurate). Such a case is discussed in Example 5.1, at the very beginning of Chapter 5. △

6.3 Task Assignments

Having a group with n members, different choices for assigning agents to questions concerning the premises induce different fixed assignments, which in turn yield the correct answer on the conclusion with different probability. In this section we are interested in finding the optimal (viz., the most accurate) such assignment. The basic trade-off that needs to be resolved in this setting is the following:

> Should we assign a large number of questions to the agents in order to collect more information, but with the implication of lowering their accuracy, or should we aim at gathering less information with higher accuracy?

6.17. Example. Suppose you have a small group with only two agents: Bea and Cal. Their accuracy is $\mathtt{q} = 0.99$ when they handle a single task, and $\mathtt{p} = 0.6$ when they multitask. Your goal is to figure out the correct answer on $\varphi \wedge \psi$ with as high a probability as possible. Is it better to ask Bea to evaluate premise φ and Cal to evaluate premise ψ, or to ask both of them to evaluate φ and ψ at the same time? You may feel inclined to pursue the former option, since then each agent will give you an extremely accurate answer on their assigned premise—this will indeed be the best option. What would you do if it instead were the case that $\mathtt{q} = 0.61$? You may then prefer to sacrifice a tiny bit of accuracy in order to get a second opinion on every premise—although this choice is intuitive, we will soon show that actually it is not the best one. △

Let us denote by $n_1 \leqslant \lfloor \frac{n}{2} \rfloor$ the number of agents that will be asked to report a judgment only on premise φ. For symmetry reasons, we assume that the same number of agents will be asked to report a judgment only on premise ψ, and the remaining $n - 2n_1$ agents will be asked to report a judgment on both premises. Given n_1, $P_{\varphi,n_1}(F_{opt})$ is the probability of the aggregation rule F_{opt} producing a correct evaluation of premise φ, and since we assume that the same number of agents that will judge φ will also judge ψ, it will be the case that $P_{\varphi,n_1}(F_{opt}) = P_{\psi,n_1}(F_{opt})$. As in Section 6.1, the accuracy of F_{opt} regarding the conclusion is $P_{n_1}(F_{opt}) = \frac{1}{2} + \frac{P_{\varphi,n_1}(F_{opt})P_{\psi,n_1}(F_{opt})}{2} = \frac{1}{2} + \frac{P_{\varphi,n_1}(F_{opt})^2}{2}$. Thus, we will maximise $P_{n_1}(F_{opt})$ if and only if we maximise $P_{\varphi,n_1}(F_{opt})$:

$$\operatorname*{argmax}_{0 \leqslant n_1 \leqslant \lfloor \frac{n}{2} \rfloor} P_{n_1}(F_{opt}) \quad = \quad \operatorname*{argmax}_{0 \leqslant n_1 \leqslant \lfloor \frac{n}{2} \rfloor} P_{\varphi,n_1}(F_{opt})$$

The optimal assignment depends on the specific number of agents n, but also on the values $\mathtt{p}$ and $\mathtt{q}$ of the individual accuracy. For small groups of at most four agents, we calculate exactly what the optimal assignment is for any $\mathtt{p}$ and $\mathtt{q}$; for larger groups, we provide results for several indicative values of $\mathtt{p}$ and $\mathtt{q}$.

6.18. Proposition. *For $n = 2$, $\operatorname{argmax}_{n_1} P_{\varphi,n_1}(F_{opt}) = 1$. Thus, when there are just two agents, it is optimal to ask each of them to evaluate one of the two premises ($n_1 = 1$) rather than asking both to evaluate both premises ($n_1 = 0$).*

Proof. For $n = 2$ we have two options: $n_1 = 0$ or $n_1 = 1$. We consider them separately. It is the case that $P_{\varphi,0}(F_{opt}) = \mathtt{p}^2 + \frac{1}{2}2\mathtt{p}(1 - \mathtt{p}) = \mathtt{p}$, while $P_{\varphi,1}(F_{opt}) = \mathtt{q} > \mathtt{p}$. Thus, $\operatorname{argmax}_{n_1} P_{\varphi,n_1}(F_{opt}) = 1$. □

6.19. Proposition. *For $n = 3$, $\operatorname{argmax}_{n_1} P_{\varphi,n_1}(F_{opt}) = 1$ if and only if $\mathtt{q} \geqslant \mathtt{p}^2(3 - 2\mathtt{p})$.*

Proof. For $n = 3$ we have two options: $n_1 = 0$ or $n_1 = 1$. We consider them separately. It is the case that $P_{\varphi,0}(F_{opt}) = \mathtt{p}^3 + \binom{3}{2}\mathtt{p}^2(1 - \mathtt{p}) = \mathtt{p}^2(3 - 2\mathtt{p})$, while $P_{\varphi,1}(F_{opt}) = \mathtt{q}$ (because the judgment of the agent who reports only on premise φ will always prevail over the judgment of the agent who reports on both premises). Thus, $\operatorname{argmax}_{n_1} P_{\varphi,n_1}(F_{opt}) = 1$ if and only if $\mathtt{q} \geqslant \mathtt{p}^2(3 - 2\mathtt{p})$. □

Thus, if agents who evaluate both premises are correct 60% of the time, then in case there are three agents, you should ask two of them to focus on a single premise each if and only if their accuracy for doing so is at least 64.8%.

6.20. Proposition. *For $n = 4$,*

$$\operatorname*{argmax}_{n_1} P_{\varphi,n_1}(F_{opt}) = \begin{cases} 1 & \textit{if } \mathsf{q} < \frac{\mathsf{p}^2}{(1-\mathsf{p})^2+\mathsf{p}^2} \\ 2 & \textit{otherwise} \end{cases}$$

Proof. We have three options: $n_1 = 0$, $n_1 = 1$, or $n_1 = 2$.

$$P_{\varphi,0}(F_{opt}) = \mathsf{p}^4 + \binom{4}{3}\mathsf{p}^3(1-\mathsf{p}) + \frac{1}{2}\binom{4}{2}\mathsf{p}^2(1-\mathsf{p})^2 = \mathsf{p}^2(3-2\mathsf{p})$$

$$P_{\varphi,1}(F_{opt}) = \begin{cases} \mathsf{q}(\mathsf{p}^2 + 2\mathsf{p}(1-\mathsf{p}) + \frac{(1-\mathsf{p})^2}{2}) + (1-\mathsf{q})\frac{\mathsf{p}^2}{2} & \text{if } \mathsf{q} = \frac{\mathsf{p}^2}{(1-\mathsf{p})^2+\mathsf{p}^2} \\ \mathsf{q} & \text{if } \mathsf{q} > \frac{\mathsf{p}^2}{(1-\mathsf{p})^2+\mathsf{p}^2} \\ \mathsf{q}(\mathsf{p}^2 + 2\mathsf{p}(1-\mathsf{p})) + (1-\mathsf{q})\mathsf{p}^2 & \text{if } \mathsf{q} < \frac{\mathsf{p}^2}{(1-\mathsf{p})^2+\mathsf{p}^2} \end{cases}$$

$$P_{\varphi,2}(F_{opt}) = \mathsf{q}^2 + 2\frac{1}{2}\mathsf{q}(1-\mathsf{q}) = \mathsf{q}$$

The claim now follows after some simple algebraic manipulations, by distinguishing cases regarding the relation of q to $\frac{\mathsf{p}^2}{(1-\mathsf{p})^2+\mathsf{p}^2}$. □

Now, for an arbitrary number of agents n and a number of agents n_1 who judge only premise φ (and the same for ψ), we have the following:

$$P_{n_1,\varphi}(F_{opt}) = \sum_{k=0}^{n-2n_1} \Big(\sum_{\substack{\ell=0 \\ \text{s.t. } \ell \in W}}^{n_1} P(k,\ell,n,n_1,\mathsf{p},\mathsf{q}) + \frac{1}{2} \sum_{\substack{\ell=0 \\ \text{s.t. } \ell \in T}}^{n_1} P(k,\ell,n,n_1,\mathsf{p},\mathsf{q}) \Big)$$

- k counts how many of the agents that judge both premises are right on φ
- ℓ counts how many of the agents that judge only φ are right on φ
- $W = \{\ell \mid \ell \log \frac{\mathsf{q}}{1-\mathsf{q}} + k \log \frac{\mathsf{p}}{1-\mathsf{p}} > (n_1 - \ell) \log \frac{\mathsf{q}}{1-\mathsf{q}} + (n - 2n_1 - k) \log \frac{\mathsf{p}}{1-\mathsf{p}}\}$
- $T = \{\ell \mid \ell \log \frac{\mathsf{q}}{1-\mathsf{q}} + k \log \frac{\mathsf{p}}{1-\mathsf{p}} = (n_1 - \ell) \log \frac{\mathsf{q}}{1-\mathsf{q}} + (n - 2n_1 - k) \log \frac{\mathsf{p}}{1-\mathsf{p}}\}$
- $P(k,\ell,n,n_1,\mathsf{p},\mathsf{q}) = \binom{n-2n_1}{k}\binom{n_1}{\ell}\mathsf{p}^k(1-\mathsf{p})^{n-2n_1-k}\mathsf{q}^\ell(1-\mathsf{q})^{n_1-\ell}$

For large groups with $n \geqslant 5$ it is too complex to calculate the optimal assignment analytically in all cases. We instead look at some representative values of p and q. For that purpose, we define a parameter α that intuitively captures the agents' *multitasking ability*, as follows: $\alpha = \frac{\mathsf{p}-0.5}{\mathsf{q}-0.5}$. Clearly, $0 < \alpha < 1$, and the smaller α is, the worse multitaskers the agents can be assumed to be.

We consider three types for the agents' multitasking ability: *good*, *average*, and *bad*, corresponding to values for α of 0.8, 0.5, and 0.2, respectively. In addition, we consider four types for the agents' accuracy on a single question: *very high*, *high*, *medium*, and *low*, corresponding to values for q of 0.9, 0.8, 0.7, and 0.6, respectively. Table 6.2 demonstrates our findings regarding the optimal assignment in terms of the number n_1 for these characteristic cases, for groups with at most 15 members.[4][5] In general, we can observe that the better multitaskers the agents are, the lower the number n_1 that corresponds to the best assignment is. This verifies an elementary intuition suggesting that if the agents are good at multitasking, then it is profitable to ask many of them about both premises, while if the agents are bad at multitasking, it is more beneficial to ask them about a single premise each.

n \ q, α	Low			Medium			High			Very high		
	bad	avg.	good	bad	avg.	good	bad	avg.	good	bad	avg.	good
5	2	2	2	2	2	2	2	2	2	2	2	2
6	3	3	3	3	3	3	3	3	3	3	3	3
7	3	3	2	3	3	2	3	3	3	3	3	3
8	4	4	3	4	3	3	4	4	3	4	4	3
9	4	4	4	4	4	4	4	4	4	4	4	4
10	5	5	5	5	5	5	5	5	5	5	5	5
11	5	5	4	5	5	4	5	5	5	5	5	5
12	6	6	5	6	5	5	6	6	5	6	6	5
13	6	6	6	6	6	6	6	6	6	6	6	6
14	7	7	7	7	7	7	7	7	7	7	7	7
15	7	7	6	7	7	6	7	7	7	7	7	7

Table 6.2: Optimal assignment in terms of the number of agents who should be asked about premise φ only, for different group size, individual accuracy, multitasking ability.

To conclude our formal analysis of truth-tracking aggregation rules under the presence of agents that multitask, let us make a few additional comments. For this first study on the topic, a few simplifying assumptions have been made. First, we have assumed that all agents have the same accuracy and that there are only two premises. Our analysis can be naturally extended beyond this case, considering different accuracies and more than two premises: The optimal aggregation rule would still be a weighted majority rule, agents with truth-oriented preferences would still always be sincere, while agents who care about their reputation or their individual opinion would still find reasons to lie—but careful further work is essential here in order to clarify all relevant details. Second, we have assumed that the exact values of the agents' accuracies are known, but this is often not true in practice. Thus, to complement our theoretical work, our results could be combined with existing experimental research that measures the accuracy of individual agents on specific application domains, ranging from human-computer interaction (Adler and Benbunan-Fich, 2012) to crowdsourcing (Cohensius et al., 2018).

[4]For *any* assignment, collective accuracy converges to 1 as the size of the group grows larger. Thus, our analysis is most interesting for groups that are not very large.

[5]The calculations were performed using a computer program in R.

Chapter 7

Conclusion

The running theme of this thesis has been *incompleteness* in collective decision making. We have delved into two formal frameworks, judgment aggregation and preference aggregation, and have provided a systematic study of a plethora of challenges that incomplete opinions bring about. Looking back, we have perhaps discovered more questions than answers—but one could also claim that this is an inherent characteristic of any attempt at scientific research. Let us recall the most notable ones amongst the questions that we did manage to answer in Section 7.1, before proceeding to a discussion of open directions for future work in Section 7.2.

7.1 Recollections

After establishing the formal background for our work in **Chapter 2**, we started our exploration in **Chapter 3**, where we inquired about natural generalisations—for the incomplete case—of aggregation rules that are widely known and respected under the assumption of complete individual opinions. Regarding the first of them, the Kemeny rule, we have proposed a way to extend it into a weighted version in which weights are assigned to the members of a collective decision process according to the degree of incompleteness that their opinions display. The various possibilities for defining such *weight rules* belong to a spectrum: at the one extreme, we have weights that are constant; at the other extreme, we can find weights that radically change under the smallest variation in an opinion's incompleteness (such as weights inducing a rule that prioritises small opinions in a lexicographic manner).

The next class of rules that we have examined is that of *quota rules* in judgment aggregation. We have introduced a taxonomy of their possible generalisations for incomplete judgments, based on two considerations: (i) whether the number of the agents that abstain should matter for the collective decision (dividing rules into variable and invariable ones), and (ii) whether we care about the absolute number of agents that support a given proposition or the margin between the agents that are in favour of the proposition and those that are against it (dividing rules into absolute and marginal ones). We have shown that variable absolute and variable marginal quota rules give rise to exactly the same class, which includes all invariable rules. But no invariable absolute rule can ever coincide with an invariable marginal rule, unless it is trivial.

In the context of preference aggregation and voting, we have focused on the popular *positional scoring rules* and have provided a general definition for this class under incompleteness. We have specifically presented two ways for extending k-approval rules, accounting for either the alternatives that a given alternative dominates or the ones by which it is dominated. We have also studied three rules that generalise the classical Borda rule in a pessimistic, optimistic, and averaging fashion, by inspecting their formal links to domination relationships between the alternatives in a given preference.

We continued in **Chapter 4**, where we searched for normative conditions (viz., *axioms*) that can be imposed on aggregation rules for incomplete individual opinions in order to guarantee reasonable and compelling collective decisions.

Within the class of weight rules, different normative concerns, encoded via a number of axioms, have been found to characterise different rules. For example, adhering to the opinion of a majority—when that opinion does not cause any problematic cases of logical inconsistency or of cyclic preferences—is achieved (only) by constant weight rules; ensuring that agents with compatible opinions will not be able to change the outcome by asserting their opinions collectively instead of individually enforces weights that are inversely proportional to the number of issues addressed by an agent.

The most general type of quota rules for incomplete judgments, i.e., variable quota rules, have been characterised by three basic axioms—*anonymity*, a weak variant of *monotonicity*, and a weak variant of *independence*—capturing the fact that quota rules only rely on the number of the votes (and not on the names of the voters), welcome additional support for the acceptance of a proposition, and perform issue-by-issue aggregation. Stronger variants of monotonicity and independence have led us to invariable absolute quota rules, while adding a *cancellation* property (that balances out positive and negative judgments on a given issue) has brought us to the class of invariable marginal quota rules. We have also hinted at the conditions under which quota rules yield logically consistent collective judgments, and have found that incompleteness creates further obstacles in that respect.

Along the lines of existing work, we have offered an axiomatic characterisation of the class of positional scoring rules in voting using *anonymity*, *neutrality*, *reinforcement*, and *continuity*. In addition, we have established that by including *monotonicity* and a variant of *cancellation* (prescribing that every alternative should be regarded a winner when the group provides equal support for all pairwise rankings), we obtain the Borda class—containing the pessimistic, the optimistic, and the averaged Borda rule. Each one of these three rules has been uniquely associated with its own axiom.

In **Chapter 5**, we turned our attention to *strategic* agents who may lie about their sincere opinions in order to achieve a collective outcome that is more attractive to them. We have proven that the rules that are immune to this kind of behaviour within judgment aggregation are exactly those that are monotonic and independent (the latter in a very weak sense). This means that amongst all quota rules for incomplete judgments, only invariable absolute ones pass the test.

Although we have not analogously characterised voting rules that are not manipulable, we have identified an exhaustive list of ways by which the agents may misrepresent their preferences when incompleteness is enabled: namely by *omitting*, *adding*, or *flipping* some of the pairwise comparisons with which their sincere preference agrees. If we only try to escape one single type of lying from the aforementioned ones, then we have proven that we can choose an appropriate positional scoring rule that achieves our goal—the same holds for the combination of addition and flipping, but fails for any combination of types that includes omission.

Then, specifically examining rules that generalise k-approval in voting with complete preferences, we have developed a model where the agents engage in *iterative* strategic behaviour, changing the collective outcome in rounds. We have shown that often there is no guarantee of arriving at stable profiles, where no agent has an incentive to further alter her preference, unless only acts of omission or only acts of addition are allowed. Complementing our findings that are based on a worst-case analysis, we have also conducted experiments suggesting that in practice stable states will always be reached, relatively quickly. Our simulations also indicate that manipulable profiles, although theoretically unavoidable, are rare.

We wrapped up our quest in **Chapter 6**, where we took a shift in perspective and studied agents whose opinions pertain to *objective* issues, with an existing—yet unknown—true answer, rather than to subjective ones. A special weight rule has been demonstrated to constitute the aggregation process that discovers the relevant truth in a probabilistically optimal manner. We have examined the strategic behaviour of agents in a group when this rule is applied, and have found that it varies vastly with respect to the preferences that the agents are assumed to hold. For example, when these preferences are in unison with the aim of the collective decision process—that is, locating the ground truth—then no incentives to lie arise. On the contrary, agents have strong incentives towards insincerity when they are aware of the opinions reported by their peers and they merely want to agree with them.

Our final question concerned the optimal method for distributing tasks to the agents in a team: By asking a certain agent to assess multiple issues simultaneously, we will gather more information, but that agent's need to multitask will lower her accuracy; on the other hand, by assigning a single task to each agent, we will collect a smaller number of more accurate pieces of evidence about the issues at stake. We have resolved this trade-off between judgment *quantity* and judgment *quality* for two, three, and four agents, and have provided numerical results for larger groups.

7.2 Reflections

By digging deeper into the foundations of incomplete opinions in the context of collective decision making, we have spotted multiple open directions for further research, which vary in character: they are technically, behaviourally, or philosophically oriented. Let us discuss them in this order.

Technical Open Questions. In many parts of this thesis we have encountered concrete technical obstacles, which pose limitations to our results (for example, no full characterisations have been found for the class of weight rules, nor for the class of rules that are immune to strategic behaviour under different kinds of manipulation moves). Since we have already addressed this kind of open questions within the relevant chapters, we now wish to have a look at the broader picture.

We should note that all aggregation rules that we have considered in the voting framework, determining a single winning alternative, constitute some kind of positional scoring rule—as we saw, this is also the case for the single-winner weight rules, which reduce to the Borda rule under complete preferences. Our study could be extended to different kinds of voting rules, such as rules that are defined based on the weighted majority graph, or approval-based rules. From a technical perspective, it would be interesting to reveal the effects of incompleteness regarding both the axiomatic and the manipulability properties of a wider pool of rules.

Moreover, techniques from preference aggregation (more abstractly, from rank aggregation) are applied to many domains besides voting, including, for example, artificial intelligence and biology. Although in this thesis we have restricted our attention to traditional aggregation rules for voting and have conducted purely theoretical work, our ideas could be transferred to applications. In particular, the axiomatic method is already employed in artificial intelligence research (e.g., by Pennock et al. (2000) with respect to collaborative filtering, which is relevant for recommender systems, and by Altman and Tennenholtz (2005) for PageRank, which is relevant for search engines)—incorporating our insights on incompleteness could generate intriguing new results.

Finally, note that we have exclusively focused on incomplete opinions when these stem from the side of the individuals in a group. But collective decisions often are incomplete as well, for instance because of ties that cannot be broken straightforwardly or because the group has to choose between a complete and inconsistent or an incomplete and consistent outcome. Incomplete collective opinions require a different treatment than individual ones. One possible way to explore this topic is through perpetual collective decision making, in which issues that have not been resolved by the group are delegated to a future decision problem.

Behavioural Open Questions. A research area that is developed in parallel to social choice theory (where this thesis belongs) revolves around the behavioural attitudes of people in settings of collective decision making. Relevant questions asked by empirical scientists concern, for example, the way that people vote under the implementation of different voting systems, or the degree to which voters' expectations about the outcome of an election affect their motivation to participate in it. An open, empirical question connected to our work is related to our investigations on truth-tracking. Specifically, even if it is clear that multitasking is part of our everyday life and that it commonly harms our performance, there is no much evidence about how people that manage teams of multitaskers decide to assign tasks to them. In addition, it would be interesting to study the final decisions that such managers make after being exposed to the judgments of their team members: are they consistent with the ones produced by our optimal aggregation rule? If it is discovered that (normative) theory and practice are largely diverting, we may want to consider better educating people about collective decision making (or perhaps, modifying our models towards a more descriptive stance). More generally, further experiments clarifying the differences between the various ways with which the agents may express the incompleteness of their opinions (e.g., by abstaining, by submitting a blank, or an invalid vote) could be very relevant for future work on collective decision making.

Philosophical Open Questions. As is the case for every endeavour to formally study phenomena where human beings are involved, the contributions of this thesis are also contingent on a number of strong assumptions about the human mind. We have built our models over a notion of incompleteness according to which the agents "may not be able to compare some of the alternatives in question". Philosophically speaking, this is a rather vague description—it does not tell us anything about where exactly this incompleteness comes from, or how it connects to the agents' beliefs and goals. Although we have not crossed the borders into the (complex, yet beautiful) philosophical terrain, we have tried to be as explicit as possible about the assumptions in our results, so that a further analysis of them in the future will be possible.

Specifically, incomplete individual opinions are enthralling in decision-theoretical terms, where one approach is to think of preference as choice: An agent choosing alternative a and not alternative b when both options are available means that the agent prefers a over b. But an agent can always be forced to choose between two options. What does this then imply about her possibly incomplete preferences? Alternatively, an agent preferring a to b may be construed as the agent expecting that a will make her happier than b (this is an interpretation of preference as welfare). In such a case, is it possible for the agent's preferences to be non-transitive (if a makes me happier than b and b makes me happier than c, then naturally a makes me happier than c)? Questions like the above (inspired by Mandler, 2005), together with many more, are profoundly tackled in philosophical decision research, but further links to formal frameworks of collective decision making can be envisioned.

Bibliography

Ackerman, M., Choi, S.-Y., Coughlin, P., Gottlieb, E., and Wood, J. (2013). Elections with partially ordered preferences. *Public Choice*, 157(1-2):145–168. (Cited on page 50.)

Adler, R. F. and Benbunan-Fich, R. (2012). Juggling on a high wire: Multitasking effects on performance. *International Journal of Human-Computer Studies*, 70(2):156–168. (Cited on pages 149 and 164.)

Aguiar-Conraria, L. and Magalhães, P. C. (2010). Referendum design, quorum rules and turnout. *Public Choice*, 144(1-2):63–81. (Cited on page 105.)

Ahn, D. S. and Oliveros, S. (2014). The Condorcet jur(ies) theorem. *Journal of Economic Theory*, 150:841–851. (Cited on page 151.)

Alcalde-Unzu, J. and Vorsatz, M. (2009). Size approval voting. *Journal of Economic Theory*, 144(3):1187–1210. (Cited on page 22.)

Altman, A. and Tennenholtz, M. (2005). Ranking systems: The pagerank axioms. In *Proceedings of the 6th ACM Conference on Electronic Commerce (EC)*. (Cited on page 168.)

Ariely, D. and Zakay, D. (2001). A timely account of the role of duration in decision making. *Acta Psychologica*, 108(2):187–207. (Cited on page 149.)

Arrow, K., Sen, A., and Suzumura, K., editors (2002). *Handbook of Social Choice and Welfare*. North-Holland. (Cited on page 1.)

Barberà, S., Bossert, W., and Pattanaik, P. K. (2004). Ranking sets of objects. In *Handbook of Utility Theory*, pages 893–977. Springer. (Cited on page 112.)

Baumeister, D., Erdélyi, G., Erdélyi, O. J., and Rothe, J. (2013). Computational aspects of manipulation and control in judgment aggregation. In *Proceedings of the 3rd*

International Conference on Algorithmic Decision Theory (ADT). Springer. (Cited on page 105.)

Baumeister, D., Faliszewski, P., Lang, J., and Rothe, J. (2012). Campaigns for lazy voters: Truncated ballots. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. (Cited on pages 49 and 50.)

Baumeister, D., Rothe, J., and Selker, A.-K. (2017). Strategic behavior in judgment aggregation. In Endriss, U., editor, *Trends in Computational Social Choice*, chapter 8, pages 145–168. AI Access. (Cited on page 106.)

Betzler, N. and Dorn, B. (2010). Towards a dichotomy for the possible winner problem in elections based on scoring rules. *Journal of Computer and System Sciences*, 76(8):812–836. (Cited on page 32.)

de Borda, J.-C. (1784). Memoire sur les elections au scrutin. *Histoire de l'Academie Royale des Sciences*. (Cited on page 16.)

Botan, S. and Endriss, U. (2020). Majority-strategyproofness in judgment aggregation. In *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. (Cited on page 105.)

Botan, S., Novaro, A., and Endriss, U. (2016). Group manipulation in judgment aggregation. In *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. (Cited on page 76.)

Bovens, L. and Rabinowicz, W. (2006). Democratic answers to complex questions—an epistemic perspective. *Synthese*, 150(1):131–153. (Cited on page 151.)

Bozbay, I., Dietrich, F., and Peters, H. (2014). Judgment aggregation in search for the truth. *Games and Economic Behavior*, 87:571–590. (Cited on page 151.)

Brandl, F. and Peters, D. (2019). An axiomatic characterization of the Borda mean rule. *Social Choice and Welfare*, 52(4):685–707. (Cited on page 88.)

Brandt, F., Conitzer, V., Endriss, U., Lang, J., and Procaccia, A. D., editors (2016). *Handbook of Computational Social Choice*. Cambridge University Press. (Cited on page 2.)

Brânzei, S., Caragiannis, I., Morgenstern, J., and Procaccia, A. D. (2013). How bad is selfish voting? In *Proceedings of the 27th AAAI Conference on Artificial Intelligence (AAAI)*. (Cited on pages 123 and 140.)

Caragiannis, I., Krimpas, G., and Voudouris, A. (2015). Aggregating partial rankings with applications to peer grading in massive online open courses. In *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. (Cited on pages 49 and 50.)

Ching, S. and Zhou, L. (2002). Multi-valued strategy-proof social choice rules. *Social Choice and Welfare*, 19(3):569–580. (Cited on page 111.)

de Clippel, G. and Eliaz, K. (2015). Premise-based versus outcome-based information aggregation. *Games and Economic Behavior*, 89:34–42. (Cited on page 151.)

Cohensius, G., Porat, O. B., Meir, R., and Amir, O. (2018). Efficient crowdsourcing via proxy voting. In *Proceedings of the 7th International Workshop on Computational Social Choice (COMSOC)*. (Cited on page 164.)

de Condorcet, N. (1785). Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix. (Cited on pages 5, 65, 149, and 151.)

Côrte-Real, P. P. and Pereira, P. T. (2004). The voter who wasn't there: Referenda, representation and abstention. *Social Choice and Welfare*, 22(2):349–369. (Cited on page 72.)

Cullinan, J., Hsiao, S. K., and Polett, D. (2014). A Borda count for partially ordered ballots. *Social Choice and Welfare*, 42(4):913–926. (Cited on pages 49 and 88.)

Dawid, A. P. and Skene, A. M. (1979). Maximum likelihood estimation of observer error-rates using the EM algorithm. *Applied Statistics*, 28(1):20–28. (Cited on page 153.)

Dery, L., Obraztsova, S., Rabinovich, Z., and Kalech, M. (2019). Lie on the fly: Strategic voting in an iterative preference elicitation process. *Group Decision and Negotiation*, 28(6):1077–1107. (Cited on page 123.)

Dietrich, F. (2014). Scoring rules for judgment aggregation. *Social Choice and Welfare*, 42(4):873–911. (Cited on pages 20 and 21.)

Dietrich, F. and List, C. (2007a). Arrow's theorem in judgment aggregation. *Social Choice and Welfare*, 29(1):19–33. (Cited on pages 13 and 59.)

Dietrich, F. and List, C. (2007b). Judgment aggregation by quota rules: Majority voting generalized. *Journal of Theoretical Politics*, 19(4):391–424. (Cited on pages 72, 74, and 76.)

Dietrich, F. and List, C. (2007c). Strategy-proof judgment aggregation. *Economics and Philosophy*, 23(03):269–300. (Cited on pages 59, 76, 105, 106, and 108.)

Dietrich, F. and List, C. (2008). Judgment aggregation without full rationality. *Social Choice and Welfare*, 31(1):15–39. (Cited on pages 4 and 75.)

Dietrich, F. and List, C. (2010a). Majority voting on restricted domains. *Journal of Economic Theory*, 145(2):512–543. (Cited on page 72.)

Dietrich, F. and List, C. (2010b). The problem of constrained judgment aggregation. In Stadler, F., editor, *The Present Situation in the Philosophy of Science*, pages 125–139. Springer. (Cited on page 59.)

Dokow, E. and Holzman, R. (2010). Aggregation of binary evaluations with abstentions. *Journal of Economic Theory*, 145(2):544–561. (Cited on page 4.)

Douceur, J. R. (2002). The Sybil attack. In *Proceedings of the 1st International Workshop on Peer-to-Peer Systems (IPTPS)*. (Cited on page 113.)

Duddy, C., Piggins, A., and Zwicker, W. S. (2016). Aggregation of binary evaluations: A Borda-like approach. *Social Choice and Welfare*, 46(2):301–333. (Cited on pages 50 and 88.)

Duggan, J. and Schwartz, T. (2000). Strategic manipulability without resoluteness or shared beliefs: Gibbard-Satterthwaite generalized. *Social Choice and Welfare*, 17(1):85–93. (Cited on page 111.)

Dummett, M. (1997). *Principles of Electoral Reform*. Oxford University Press. (Cited on pages 49 and 50.)

Edland, A. and Svenson, O. (1993). Judgment and decision making under time pressure. In *Time Pressure and Stress in Human Judgment and Decision Making*, pages 27–40. Springer. (Cited on page 149.)

Emerson, P. (2013). The original Borda count and partial voting. *Social Choice and Welfare*, 40(2):353–358. (Cited on page 49.)

Endriss, U. (2016). Judgment aggregation. In Brandt, F., Conitzer, V., Endriss, U., Lang, J., and Procaccia, A. D., editors, *Handbook of Computational Social Choice*, pages 399–426. Cambridge University Press. (Cited on page 10.)

Endriss, U. (2018). Judgment aggregation with rationality and feasibility constraints. In *Proceedings of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. (Cited on page 31.)

Endriss, U., Obraztsova, S., Polukarov, M., and Rosenschein, J. S. (2016). Strategic voting with incomplete information. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI)*. (Cited on page 123.)

Endriss, U., Pini, M. S., Rossi, F., and Venable, K. B. (2009). Preference aggregation over restricted ballot languages: Sincerity and strategy-proofness. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI)*. (Cited on page 29.)

Faliszewski, P., Hemaspaandra, E., and Hemaspaandra, L. (2010). Using complexity to protect elections. *Communications of the ACM*, 53(11):74–82. (Cited on page 2.)

Faliszewski, P., Skowron, P., Slinko, A., and Talmon, N. (2017). Multiwinner voting: A new challenge for social choice theory. In *Trends in Computational Social Choice*. AI Access. (Cited on pages 17 and 28.)

Farkas, D. and Nitzan, S. (1979). The Borda rule and Pareto stability: A comment. *Econometrica*, 47(5):1305–1306. (Cited on page 88.)

Fishburn, P. C. (1970). Intransitive individual indifference and transitive majorities. *Econometrica*, 38(3):482–489. (Cited on page 5.)

Fishburn, P. C. (1973). *The Theory of Social Choice*. Princeton University Press. (Cited on page 73.)

Fishburn, P. C. and Brams, S. J. (1983). Paradoxes of preferential voting. *Mathematics Magazine*, 56(4):207–214. (Cited on page 105.)

Fishburn, P. C. and Gehrlein, W. V. (1976). Borda's rule, positional voting, and Condorcet's simple majority principle. *Public Choice*, 28:79–88. (Cited on page 93.)

Freixas, J. and Zwicker, W. S. (2009). Anonymous yes–no voting with abstention and multiple levels of approval. *Games and Economic Behavior*, 67(2):428–444. (Cited on page 78.)

Gärdenfors, P. (1976). Manipulation of social choice functions. *Journal of Economic Theory*, 13(2):217–228. (Cited on page 111.)

Gärdenfors, P. (2006). A representation theorem for voting with logical consequences. *Economics and Philosophy*, 22(2):181–190. (Cited on pages 4, 11, and 75.)

Gehrlein, W. V. (2002). Condorcet's paradox and the likelihood of its occurrence: different perspectives on balanced preferences. *Theory and decision*, 52(2):171–199. (Cited on page 5.)

Gibbard, A. (1973). Manipulation of voting schemes: A general result. *Econometrica*, 41(4):587–601. (Cited on pages 2 and 110.)

Glasser, G. (1959). Game theory and cumulative voting for corporate directors. *Management Science*, 5(2):151–156. (Cited on pages 22 and 30.)

Grandi, U., Loreggia, A., Rossi, F., Venable, K. B., and Walsh, T. (2013). Restricted manipulation in iterative voting: Condorcet efficiency and Borda score. In *Proceedings of the 3rd International Conference on Algorithmic Decision Theory (ADT)*. (Cited on pages 123, 141, and 142.)

Grossi, D. and Pigozzi, G. (2014). *Judgment Aggregation: A Primer*. Synthesis Lectures on Artificial Intelligence and Machine Learning. Morgan and Claypool Publishers. (Cited on page 10.)

Hansson, B. and Sahlquist, H. (1976). A proof technique for social choice with variable electorate. *Journal of Economic Theory*, 13(2):1027–1041. (Cited on page 99.)

Hartmann, S., Pigozzi, G., and Sprenger, J. (2010). Reliable methods of judgement aggregation. *Journal of Logic and Computation*, 20(2):603–617. (Cited on page 151.)

Hartmann, S. and Sprenger, J. (2012). Judgment aggregation and the problem of tracking the truth. *Synthese*, 187(1):209–221. (Cited on page 151.)

Houy, N. (2007). Some further characterizations for the forgotten voting rules. *Mathematical Social Sciences*, 53(1):111–121. (Cited on pages 73, 74, and 78.)

Jiang, G., Zhang, D., and Perrussel, L. (2018). A hierarchical approach to judgment aggregation with abstentions. *Computational Intelligence*, 34(1):104–123. (Cited on page 4.)

Kelly, J. S. (1977). Strategy-proofness and social choice functions without singlevaluedness. *Econometrica*, 45(2):439–446. (Cited on page 111.)

Kelly, J. S. (1993). Almost all social choice rules are highly manipulable, but a few aren't. *Social Choice and Welfare*, 10(2):161–175. (Cited on page 141.)

Konczak, K. and Lang, J. (2005). Voting procedures with incomplete preferences. In *Proceedings of the IJCAI Multidisciplinary Workshop on Advances in Preference Handling (MPREF)*. (Cited on pages 5 and 32.)

Koolyk, A., Lev, O., and Rosenschein, J. S. (2017). Convergence and quality of iterative voting under non-scoring rules. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI)*. (Cited on page 123.)

Kornhauser, L. A. and Sager, L. G. (1993). The one and the many: Adjudication in collegial courts. *California Law Review*, 81(1):1–59. (Cited on page 55.)

Kruger, J. and Terzopoulou, Z. (2020). Strategic manipulation with incomplete preferences: Possibilities and impossibilities for positional scoring rules. In *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. (Cited on pages 6, 7, and 8.)

Lang, J., Monnot, J., Slinko, A., and Zwicker, W. S. (2017). Beyond electing and ranking: Collective dominating chains, dominating subsets and dichotomies. In *Proceedings of the 16th Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. (Cited on pages 28 and 31.)

Laslier, J.-F. (2009). The leader rule: A model of strategic approval voting in a large electorate. *Journal of Theoretical Politics*, 21(1):113–136. (Cited on page 140.)

Lev, O. and Rosenschein, J. S. (2012). Convergence of iterative voting. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. (Cited on page 123.)

Lev, O. and Rosenschein, J. S. (2016). Convergence of iterative scoring rules. *Journal of Artificial Intelligence Research*, 57:573–591. (Cited on pages 123, 131, and 133.)

List, C. (2012). The theory of judgment aggregation: An introductory review. *Synthese*, 187(1):179–207. (Cited on page 10.)

List, C. and Pettit, P. (2002). Aggregating sets of judgments: An impossibility result. *Economics and Philosophy*, 18(1):89–110. (Cited on pages 3 and 10.)

List, C. and Pettit, P. (2004). Aggregating sets of judgments: Two impossibility results compared. *Synthese*, 140(1-2):207–235. (Cited on page 13.)

Llamazares, B. (2006). The forgotten decision rules: Majority rules based on difference of votes. *Mathematical Social Sciences*, 51(3):311–326. (Cited on pages 73, 75, and 78.)

Mandler, M. (2005). Incomplete preferences and rational intransitivity of choice. *Games and Economic Behavior*, 50(2):255–277. (Cited on page 169.)

Maniquet, F. and Morelli, M. (2015). Approval quorums dominate participation quorums. *Social Choice and Welfare*, 45(1):1–27. (Cited on page 105.)

May, K. O. (1952). A set of independent necessary and sufficient conditions for simple majority decision. *Econometrica*, pages 680–684. (Cited on page 72.)

Meir, R. (2017). Iterative voting. In Endriss, U., editor, *Trends in Computational Social Choice*. AI Access. (Cited on page 122.)

Meir, R., Lev, O., and Rosenschein, J. S. (2014). A local-dominance theory of voting equilibria. In *Proceedings of the 15th ACM Conference on Economics and Computation (EC)*. (Cited on pages 123, 142, and 148.)

Meir, R., Polukarov, M., Rosenschein, J. S., and Jennings, N. R. (2010). Convergence to equilibria in plurality voting. In *Proceedings of the 24th AAAI Conference on Artificial Intelligence (AAAI)*. (Cited on pages 122, 123, 127, and 128.)

Meir, R., Polukarov, M., Rosenschein, J. S., and Jennings, N. R. (2017). Iterative voting and acyclic games. *Artificial Intelligence*, 252:100–122. (Cited on page 123.)

Miller, M. K. and Osherson, D. (2009). Methods for distance-based judgment aggregation. *Social Choice and Welfare*, 32(4):575–601. (Cited on page 12.)

Myerson, R. B. (1995). Axiomatic derivation of scoring rules without the ordering assumption. *Social Choice and Welfare*, 12(1):59–74. (Cited on pages 83, 84, and 86.)

Nehring, K., Pivato, M., and Puppe, C. (2014). The condorcet set: Majority voting over interconnected propositions. *Journal of Economic Theory*, 151:268–303. (Cited on pages 12 and 57.)

Nitzan, S. (1985). The vulnerability of point-voting schemes to preference variation and strategic manipulation. *Public choice*, 47(2):349–370. (Cited on page 141.)

Nitzan, S. and Rubinstein, A. (1981). A further characterization of Borda ranking method. *Public Choice*, 36:153–158. (Cited on page 88.)

Nurmi, H. (1992). An assessment of voting system simulations. *Public Choice*, 73(4):459–487. (Cited on page 142.)

Obraztsova, S., Lev, O., Polukarov, M., Rabinovich, Z., and Rosenschein, J. S. (2015a). Farsighted voting dynamics. In *Proceedings of the 1st IJCAI Workshop on Algorithmic Decision Theory*. (Cited on page 123.)

Obraztsova, S., Markakis, E., Polukarov, M., Rabinovich, Z., and Jennings, N. R. (2015b). On the convergence of iterative voting: How restrictive should restricted dynamics be? In *Proceedings of the 29th AAAI Conference on Artificial Intelligence (AAAI)*. (Cited on page 124.)

Obraztsova, S., Markakis, E., and Thompson, D. R. (2013). Plurality voting with truth-biased agents. In *Proceedings of the 6th International Symposium on Algorithmic Game Theory (SAGT)*. (Cited on page 123.)

Payne, J. W., Bettman, J. R., and Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14(3):534. (Cited on page 149.)

Pennock, D. M., Horvitz, E., Giles, C. L., et al. (2000). Social choice theory and recommender systems: Analysis of the axiomatic foundations of collaborative filtering. In *Proceedings of the 17th AAAI Conference on Artificial Intelligence (AAAI)*. (Cited on page 168.)

Pettit, P. (2001). Deliberative democracy and the discursive dilemma. *Noûs*, 35:268–299. (Cited on page 3.)

Pigozzi, G. (2006). Belief merging and the discursive dilemma: An argument-based account to paradoxes of judgment aggregation. *Synthese*, 152(2):285–298. (Cited on page 12.)

Pini, M. S., Rossi, F., Venable, K. B., and Walsh, T. (2007). Incompleteness and incomparability in preference aggregation. In *Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI)*. (Cited on pages 5 and 32.)

Pini, M. S., Rossi, F., Venable, K. B., and Walsh, T. (2008). Aggregating partially ordered preferences. *Journal of Logic and Computation*, 19(3):475–502. (Cited on pages 5 and 111.)

Pini, M. S., Rossi, F., Venable, K. B., and Walsh, T. (2011). Incompleteness and incomparability in preference aggregation: Complexity results. *Artificial Intelligence*, 175(7-8):1272–1289. (Cited on page 5.)

Rabinovich, Z., Obraztsova, S., Lev, O., Markakis, E., and Rosenschein, J. S. (2015). Analysis of equilibria in iterative voting schemes. In *Proceedings of the 29th AAAI Conference on Artificial Intelligence (AAAI)*. (Cited on page 123.)

Reijngoud, A. and Endriss, U. (2012). Voter response to iterated poll information. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. (Cited on pages 123, 125, and 142.)

Reyhani, R. and Wilson, M. C. (2012). Best reply dynamics for scoring rules. In *Proceedings of the 20th European Conference on Artificial Intelligence (ECAI)*. (Cited on pages 123 and 128.)

Saari, D. G. (1990). The Borda dictionary. *Social Choice and Welfare*, 7(4):279–317. (Cited on page 88.)

Saari, D. G. (2008). *Disposing Dictators, Demystifying Voting Paradoxes*. Cambridge University Press. (Cited on pages 49 and 50.)

Sato, S. (2008). On strategy-proof social choice correspondences. *Social Choice and Welfare*, 31(2):331–343. (Cited on page 111.)

Satterthwaite, M. A. (1975). Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory*, 10(2):187–217. (Cited on pages 2 and 110.)

Schechter, E. (1996). *Handbook of Analysis and its Foundations*. Academic Press. (Cited on page 25.)

Slavkovik, M. and Jamroga, W. (2011). Distance-based judgment aggregation of three-valued judgments with weights. In *Proceedings of the IJCAI Workshop on Social Choice and Artificial Intelligence*. (Cited on pages 4 and 22.)

Smith, J. H. (1973). Aggregation of preferences with variable electorate. *Econometrica*, 41(6):1027–1041. (Cited on pages 82, 83, 93, and 94.)

Terzopoulou, Z. (2020). Quota rules for incomplete judgments. *Mathematical Social Sciences*, 107:23–36. (Cited on pages 6, 7, 8, and 42.)

Terzopoulou, Z. and Endriss, U. (2018). Modelling iterative judgment aggregation. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence (AAAI)*. (Cited on page 140.)

Terzopoulou, Z. and Endriss, U. (2019a). Aggregating incomplete pairwise preferences by weight. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI)*. (Cited on pages 6, 7, and 29.)

Terzopoulou, Z. and Endriss, U. (2019b). Optimal truth-tracking rules for the aggregation of incomplete judgments. In *Proceedings of the 12th International Symposium on Algorithmic Game Theory (SAGT)*. (Cited on page 8.)

Terzopoulou, Z. and Endriss, U. (2019c). Strategyproof judgment aggregation under partial information. *Social Choice and Welfare*, 53(3):415–442. (Cited on page 105.)

Terzopoulou, Z. and Endriss, U. (2020). Neutrality and relative acceptability in judgment aggregation. *Social Choice and Welfare*, 55(1):25–49. (Cited on page 76.)

Terzopoulou, Z. and Endriss, U. (2021). The Borda class: An axiomatic study of the Borda rule on top-truncated preferences. *Journal of Mathematical Economics*, 92:31–40. (Cited on pages 6 and 7.)

Terzopoulou, Z., Endriss, U., and de Haan, R. (2018). Aggregating incomplete judgments: Axiomatisations for scoring rules. In *Proceedings of the 7th International Workshop on Computational Social Choice (COMSOC)*. (Cited on pages 6 and 7.)

Tsetlin, I., Regenwetter, M., and Grofman, B. (2003). The impartial culture maximizes the probability of majority cycles. *Social Choice and Welfare*, 21(3):387–398. (Cited on page 142.)

Wilhelm, O. and Schulze, R. (2002). The relation of speeded and unspeeded reasoning with mental speed. *Intelligence*, 30(6):537–554. (Cited on page 149.)

Wilson, R. (1975). On the theory of aggregation. *Journal of Economic Theory*, 10(1):89–99. (Cited on page 3.)

Xia, L. and Conitzer, V. (2008). Determining possible and necessary winners under common voting rules given partial orders. In *Proceedings of the 23rd AAAI Conference on Artificial Intelligence (AAAI)*. (Cited on pages 32 and 34.)

Xia, L. and Conitzer, V. (2011). Determining possible and necessary winners given partial orders. *Journal of Artificial Intelligence Research*, 41:25–67. (Cited on page 5.)

Yokoo, M., Sakurai, Y., and Matsubara, S. (2001). Robust combinatorial auction protocol against false-name bids. *Artificial Intelligence*, 130(2):167–181. (Cited on page 113.)

Young, H. P. (1974). An axiomatization of Borda's rule. *Journal of Economic Theory*, 9(1):43–52. (Cited on pages 88, 90, 91, 92, and 99.)

Young, H. P. (1975). Social choice scoring functions. *Journal on Applied Mathematics*, 28(4):824–838. (Cited on page 83.)

Zwicker, W. S. (2016). Introduction to the theory of voting. In *Handbook of Computational Social Choice*. Cambridge University Press. (Cited on pages 16, 28, 30, and 85.)

Zwicker, W. S. (2018). Cycles and intractability in a large class of aggregation rules. *Journal of Artificial Intelligence Research*, 61:407–431. (Cited on page 5.)

List of Symbols

$\rhd$	strict, acyclic, and transitive preference	14
$\succsim$	acyclic and transitive preference, with indifference	14
$<_\varphi$	relation between two judgment sets, encoding that the support for φ strictly increases from one to the other	74
A	set of alternatives	13
$A^\star$	set of potential alternatives	13
$\mathcal{D}(A)$	domain of all acyclic preference relations over A, with indifference	13
$\mathcal{D}(A)^t$	domain of all acyclic and transitive preference relations over A, with indifference	14
Φ	agenda	11
$\Phi^\star$	superagenda	10
$\widetilde{\varphi}$	binary issue	11
F	aggregation rule	10
F_a	invariable absolute quota rule	36
F_c	weight rule induced by the constant weight function	22
F_{dlex}	downward-lexicographic rule	25
F_{ee}	weight rule induced by the equal-and-even weight function	22
F_m	invariable marginal quota rule	36
F_{opt}	optimal rule for discovering a ground truth	152
F_{ulex}	upward-lexicographic rule	25
F_{va}	variable absolute quota rule	36
F_{vm}	variable marginal quota rule	37
F_w	weight rule	22

Index

Samenvatting

In onze moderne wereld maken mensen voortdurend beslissingen samen: men neemt deel aan verkiezingen, het bestuur van een bedrijf, de jury in een rechtszaak, en sociale activiteiten met vrienden. Het is van essentieel belang om dit soort verfijnd menselijk gedrag goed te begrijpen, niet alleen om de keuzes die mensen maken direct te verbeteren, maar ook om AI- en multiagent-systemen te ontwikkelen die op een efficiënte en veilige manier samenwerken met collectieve beslissingen. Socialekeuzetheorie is het deelgebied van de micro-economie die zich bezighoudt met het analyseren van mechanismen die groepen gebruiken om beslissingen te maken; de thema's die een rol spelen in dit onderzoeksgebied overlappen met andere onderzoeksgebieden, zoals politicologie, rechtspraak, en kunstmatige intelligentie. Het veel jongere onderzoeksgebied van computationele socialekeuzetheorie onderzoekt onderwerpen en problemen uit de socialekeuzetheorie met behulp van krachtige hulpmiddelen uit de informatica.

Veel onderzoek in computationele socialekeuzetheorie is gebaseerd op de nogal stricte aanname dat mensen een concrete mening kunnen vormen over alle kwesties die een rol spelen bij een gegeven beslissing (zulke meningen noemen we volledig). In de afgelopen jaren maken mensen echter steeds vaker beslissingen op het internet, waar de overvloed aan mogelijke opties ervoor zorgt dat mensen niet alle opties kritisch kunnen evalueren. Deze dissertatie behandelt collectieve beslissingen waar de individuen inherent onvolledige meningen kunnen hebben (dat wil zeggen, het is mogelijk dat ze over een aantal relevante kwesties simpelweg geen mening hebben). Op een reiswebsite waar gebruikers hotels rangschikken aan de hand van hun ervaringen en voorkeuren, hebben de meeste gebruikers niet een mening over ieder hotel, bijvoorbeeld.

We hebben aggregatiemechanismen nodig die een onvolledige invoer toestaan en met zo'n invoer toch redelijke collectieve beslissingen produceren. Deze dissertatie biedt zulke mechanismen in drie verschillende contexten, waar we belang hebben bij (i) het bereiken van een consensus tussen de diverse meningen van de groepsleden over kwesties die subjectief van aard zijn (zoals in het voorbeeld van de reiswebsite), (ii) het vermijden van scheefgetrokken collectieve beslissingen die veroorzaakt worden door strategisch gedrag van egoïstische individuen (zoals bij verkiezingen het geval kan

zijn), en (iii) het ontdekken van de onderliggende waarheid als het gaat om een aantal objectieve kwesties (zoals bij crowdsourcing-experimenten).

Ten eerste, hoe kunnen we een adequaat compromis bereiken voor een groep individuen met uiteenlopende meningen? Eén manier om dit te doen, die vaak gebruikt wordt in de onderzoeksgemeenschap van de socialekeuzetheorie, is door gebruik te maken van axioma's: eigenschappen van aggregatiemechanismen die aantrekkelijk zijn vanwege hun normatieve interpretatie. Zulke axioma's zijn goed bestudeerd in het gebruikelijke raamwerk waar volledigheid van meningen wordt aangenomen, maar roepen talrijke open vragen op zodra onvolledigheid toegestaan wordt. Door onze aandacht te richten op verschillende axioma's verkrijgen we verschillende mechanismen voor collectieve keuzes—inclusief mechanismen die meningen van verschillende grootte niet even zwaar meewegen, mechanismen die scores toewijzen aan opties op basis van hun positie in de inviduele meningen, en mechanismen die gebaseerd zijn op verschillende acceptatiedrempels. De verschillende axioma's die we beschouwen leiden tot verschillende noties van compromis. Afhankelijk van welke men overtuigend vindt, kan geargumenteerd worden om de bijbehorende mechanismen te gebruiken.

Ten tweede, bij welke mechanismen voor groepsbeslissingen kunnen de leden van een groep een verkeerde voorstelling geven van hun individuele mening met het doel om een betere uitkomst te bereiken, vanuit hun egoïstische perspectief? Individuen die de vrijheid hebben om onvolledige meningen te rapporteren kunnen op drie manieren liegen: door hun oprechte mening te verzwijgen, door een onoprechte mening te verzinnen, of door hun oprechte mening om te draaien. Zulke daden vormen een bedreiging voor ons dagelijkse leven, doordat ze het risico met zich meebrengen op niet-representatieve, ondemocratische groepsbeslissingen. We laten zien dat het in veel gevallen mogelijk is om mechanismen te vinden die immuun zijn voor dit soort strategisch gedrag. We tonen echter ook aan dat strategische manipulatie in sommige gevallen onvermijdelijk is. Bovendien onderzoeken we, voor het geval waarin individuen hun mening mogen veranderen in een aantal rondes, of—en hoe snel—relevante iteratieve beslissingsprocedures gegarandeerd eindigen.

Ten derde, stel dat je het correcte antwoord op een vraag moet vinden (bijvoorbeeld, of een collega promotie moet krijgen) die afhangt van twee onafhankelijke premissen (bijvoorbeeld, of de collega goed is in haar werk en of ze een goede teamspeler is). Vraag je de leden van een groep om elk een van de premissen te evalueren (wat met relatief hoge nauwkeurigheid gedaan kan worden) of elk beide premissen (wat vanwege de noodzaak om te multitasken de nauwkeurigheid naar beneden zal halen)? We bieden een optimaal mechanisme om groepsbeslissingen te maken in scenario's zoals dit. We stellen vast hoeveel individuen men moet vragen om hoeveel premissen te evalueren, en we bieden een uitgebreide analyse van de strategische prikkels die zich vanuit individueel perspectief kunnen voordoen.

In een notendop samengevat: deze dissertatie bestudeert een aantal veelbelovende manieren om groepsbeslissingen te maken die toepasbaar zijn in een ruime hoeveelheid contexten (ingegeven door daadwerkelijke toepassingen) waarin individuele meningen naar verwachting onvolledig zijn.

Abstract

In the modern world, people make decisions together every day: they participate in elections, boards of companies, juries of courts, and social activities amongst friends. Accurately capturing such fine-grained human behaviour is essential, not only for directly improving the choices made by people, but also for developing AI (multiagent) systems that work efficiently and safely with collective decisions. Social choice theory, a sub-area of (micro)economics, is concerned with the analysis of mechanisms that groups use to make decisions; its far-reaching themes transcend neighbouring disciplines, like political science, law, and artificial intelligence. The much younger field of computational social choice investigates problems traditionally studied in social choice theory with powerful tools from computer science.

Much research to date in computational social choice has been resting on a rather stringent hypothesis—namely, that people are able to form concrete opinions about all issues at stake in a given decision problem (such opinions are called complete). But in recent years, people make decisions increasingly in online settings, where the abundance of available options often makes them prohibitively difficult to critically evaluate. This thesis concerns collective decisions where the agents may hold intrinsically incomplete opinions (that is, they may lack an opinion about some issues in question). For example, on a travel website where users rank hotels according to their experience and preferences, most people will not have an opinion about all offered listings.

We need aggregation mechanisms that allow for incomplete inputs and produce reasonable collective outcomes. This thesis proposes such mechanisms in three contexts, where we care about (*i*) reaching a consensus between the diverse opinions of the group members for issues that are subjective in nature (like in the example about the travel website); (*ii*) avoiding skewed outcomes caused by the strategic behaviour of selfish agents (like in election problems); (*iii*) discovering the ground truth with respect to a number of objective issues (like in crowdsourcing experiments).

First, how can we achieve a sufficiently good compromise for a group of agents with diverting opinions? One way, typically employed by the social choice community, suggests the use of axioms, i.e., properties of aggregation mechanisms that are normatively appealing. Such axioms are well-studied in the standard frameworks where completeness is assumed, but present numerous open questions when incompleteness comes into the picture. By restricting attention to different axioms, we obtain distinct mechanisms for collective choice—including mechanisms that weigh individual opinions relying on their size, mechanisms that assign scores to alternatives with respect to their position in the individual preferences, as well as mechanisms that are associated with given acceptance thresholds. The various axioms that we consider lead to different notions of compromise. Depending on those one finds compelling, one can argue for the use of the corresponding mechanisms that satisfy them.

Second, under what mechanisms for collective decision making are the members of a group tempted to misrepresent their individual opinions in view of obtaining a better outcome for themselves? Agents that have the freedom to report incomplete opinions may lie in three ways: by hiding their sincere opinion; by inventing a new insincere opinion; or by reversing their sincere opinion. Such acts pose threats to our everyday life, bringing out the risk of obtaining non-representative, undemocratic collective decisions. In many cases, we show that it is possible to find mechanisms that are immune to this kind of strategic behaviour. However, we also prove impossibility results stating that strategic manipulation is, sometimes, inescapable. Furthermore, when agents are allowed to change their opinions in rounds, we examine whether—and how fast—relevant procedures of iterative decision making are guaranteed to terminate.

Third, suppose you need to determine the correct answer to a question (e.g., whether a colleague should get promoted) that depends on two independent premises (e.g., whether the colleague is excellent at her work and whether she is a good team player). Will you ask the members of a group to each evaluate just one of those premises (which can be done with relatively high accuracy) or both (in which case their need to multitask will lower their accuracy)? In scenarios like this, we present an optimal mechanism for making collective decisions, we determine how many agents one should ask for how many of their opinions, and we provide an extensive analysis of strategic incentives that may arise from the point of view of the group members.

In a nutshell, spanning across ample contexts motivated by real-world applications, this thesis discusses several promising ways for making collective decisions when individual opinions are expected to be incomplete.

Titles in the ILLC Dissertation Series:

ILLC DS-2016-01: **Ivano A. Ciardelli**
Questions in Logic

ILLC DS-2016-02: **Zoé Christoff**
Dynamic Logics of Networks: Information Flow and the Spread of Opinion

ILLC DS-2016-03: **Fleur Leonie Bouwer**
What do we need to hear a beat? The influence of attention, musical abilities, and accents on the perception of metrical rhythm

ILLC DS-2016-04: **Johannes Marti**
Interpreting Linguistic Behavior with Possible World Models

ILLC DS-2016-05: **Phong Lê**
Learning Vector Representations for Sentences - The Recursive Deep Learning Approach

ILLC DS-2016-06: **Gideon Maillette de Buy Wenniger**
Aligning the Foundations of Hierarchical Statistical Machine Translation

ILLC DS-2016-07: **Andreas van Cranenburgh**
Rich Statistical Parsing and Literary Language

ILLC DS-2016-08: **Florian Speelman**
Position-based Quantum Cryptography and Catalytic Computation

ILLC DS-2016-09: **Teresa Piovesan**
Quantum entanglement: insights via graph parameters and conic optimization

ILLC DS-2016-10: **Paula Henk**
Nonstandard Provability for Peano Arithmetic. A Modal Perspective

ILLC DS-2017-01: **Paolo Galeazzi**
Play Without Regret

ILLC DS-2017-02: **Riccardo Pinosio**
The Logic of Kant's Temporal Continuum

ILLC DS-2017-03: **Matthijs Westera**
Exhaustivity and intonation: a unified theory

ILLC DS-2017-04: **Giovanni Cinà**
Categories for the working modal logician

ILLC DS-2017-05: **Shane Noah Steinert-Threlkeld**
Communication and Computation: New Questions About Compositionality

ILLC DS-2017-06: **Peter Hawke**
The Problem of Epistemic Relevance

ILLC DS-2017-07: **Aybüke Özgün**
Evidence in Epistemic Logic: A Topological Perspective

ILLC DS-2017-08: **Raquel Garrido Alhama**
Computational Modelling of Artificial Language Learning: Retention, Recognition & Recurrence

ILLC DS-2017-09: **Miloš Stanojević**
Permutation Forests for Modeling Word Order in Machine Translation

ILLC DS-2018-01: **Berit Janssen**
Retained or Lost in Transmission? Analyzing and Predicting Stability in Dutch Folk Songs

ILLC DS-2018-02: **Hugo Huurdeman**
Supporting the Complex Dynamics of the Information Seeking Process

ILLC DS-2018-03: **Corina Koolen**
Reading beyond the female: The relationship between perception of author gender and literary quality

ILLC DS-2018-04: **Jelle Bruineberg**
Anticipating Affordances: Intentionality in self-organizing brain-body-environment systems

ILLC DS-2018-05: **Joachim Daiber**
Typologically Robust Statistical Machine Translation: Understanding and Exploiting Differences and Similarities Between Languages in Machine Translation

ILLC DS-2018-06: **Thomas Brochhagen**
Signaling under Uncertainty

ILLC DS-2018-07: **Julian Schlöder**
Assertion and Rejection

ILLC DS-2018-08: **Srinivasan Arunachalam**
Quantum Algorithms and Learning Theory

ILLC DS-2018-09: **Hugo de Holanda Cunha Nobrega**
Games for functions: Baire classes, Weihrauch degrees, transfinite computations, and ranks

ILLC DS-2018-10: **Chenwei Shi**
Reason to Believe

ILLC DS-2018-11: **Malvin Gattinger**
New Directions in Model Checking Dynamic Epistemic Logic

ILLC DS-2018-12: **Julia Ilin**
Filtration Revisited: Lattices of Stable Non-Classical Logics

ILLC DS-2018-13: **Jeroen Zuiddam**
Algebraic complexity, asymptotic spectra and entanglement polytopes

ILLC DS-2019-01: **Carlos Vaquero**
What Makes A Performer Unique? Idiosyncrasies and commonalities in expressive music performance

ILLC DS-2019-02: **Jort Bergfeld**
Quantum logics for expressing and proving the correctness of quantum programs

ILLC DS-2019-03: **András Gilyén**
Quantum Singular Value Transformation & Its Algorithmic Applications

ILLC DS-2019-04: **Lorenzo Galeotti**
The theory of the generalised real numbers and other topics in logic

ILLC DS-2019-05: **Nadine Theiler**
Taking a unified perspective: Resolutions and highlighting in the semantics of attitudes and particles

ILLC DS-2019-06: **Peter T.S. van der Gulik**
Considerations in Evolutionary Biochemistry

ILLC DS-2019-07: **Frederik Möllerström Lauridsen**
Cuts and Completions: Algebraic aspects of structural proof theory

ILLC DS-2020-01: **Mostafa Dehghani**
Learning with Imperfect Supervision for Language Understanding

ILLC DS-2020-02: **Koen Groenland**
Quantum protocols for few-qubit devices

ILLC DS-2020-03: **Jouke Witteveen**
Parameterized Analysis of Complexity

ILLC DS-2020-04: **Joran van Apeldoorn**
A Quantum View on Convex Optimization

ILLC DS-2020-05: **Tom Bannink**
Quantum and stochastic processes

ILLC DS-2020-06: **Dieuwke Hupkes**
Hierarchy and interpretability in neural models of language processing

ILLC DS-2020-07: **Ana Lucia Vargas Sandoval**
On the Path to the Truth: Logical & Computational Aspects of Learning

ILLC DS-2020-08: **Philip Schulz**
Latent Variable Models for Machine Translation and How to Learn Them

ILLC DS-2020-09: **Jasmijn Bastings**
A Tale of Two Sequences: Interpretable and Linguistically-Informed Deep Learning for Natural Language Processing

ILLC DS-2020-10: **Arnold Kochari**
Perceiving and communicating magnitudes: Behavioral and electrophysiological studies

ILLC DS-2020-11: **Marco Del Tredici**
Linguistic Variation in Online Communities: A Computational Perspective

ILLC DS-2020-12: **Bastiaan van der Weij**
Experienced listeners: Modeling the influence of long-term musical exposure on rhythm perception

ILLC DS-2020-13: **Thom van Gessel**
Questions in Context

ILLC DS-2020-14: **Gianluca Grilletti**
Questions & Quantification: A study of first order inquisitive logic

ILLC DS-2020-15: **Tom Schoonen**
Tales of Similarity and Imagination. A modest epistemology of possibility

ILLC DS-2020-16: **Ilaria Canavotto**
Where Responsibility Takes You: Logics of Agency, Counterfactuals and Norms