



# Democracy and the Common Good

A Study of the Weighted Majority Rule

Katharina Berndt Rasmussen

©Katharina Berndt Rasmussen, Stockholm 2013

ISBN 978-91-7447-738-2

Printed in Sweden by US-AB, Stockholm 2013

Distributor: Department of Philosophy, Stockholm University

To Nika



# Contents

1 Introduction .....	11
1.1 Majority conundrums .....	11
1.2 The ubiquity of the unequal vote .....	15
1.3 The value of democracy .....	17
1.4 Main thesis and disposition .....	19
2 The basics.....	23
2.1 Introduction.....	23
2.2 The weighted majority rule .....	23
2.2.1 The weighted majority rule and democratic theory .....	25
2.3 Individual well-being and the common good .....	31
2.3.1 Individual well-being.....	31
2.3.2 The common good .....	34
2.3.3 Terminology .....	40
2.4 Democracy and the promotion of the common good.....	40
2.4.1 Arrow's theorem and related problems .....	41
2.4.2 Returning to the weighted majority rule.....	43
2.5 Notes on methodology and philosophical framework.....	45
2.5.1 The moral-philosophical framework .....	45
2.5.2 Formal models .....	46
3 A case for the weighted majority rule.....	51
3.1 Introduction.....	51
3.2 The original argument from collective optimality .....	51
3.2.1 Alternative versions of the argument.....	56
3.3 The generic argument from collective optimality .....	61
3.4 Further clarifications .....	63
3.5 Conclusions .....	70
4 Self- and common-interested voting .....	71
4.1 Introduction.....	71
4.2 The extended argument from collective optimality .....	71
4.2.1 Rebutting the 'mixed motivation' problem .....	74
4.3 Erratic voting behaviour.....	76
4.3.1 Why a voter may vote erratically .....	77
4.4 Conclusions .....	80

5	Less than fully competent voters .....	83
5.1	Introduction .....	83
5.2	The argument from weak collective optimality .....	85
5.2.1	Introducing the Condorcet jury theorem .....	85
5.2.2	Equal-stakes cases with common-interested voters .....	90
5.2.3	Equal-stakes cases with self-interested voters .....	92
5.2.4	Equal-stakes mixed-motivation cases .....	99
5.2.5	'Mixed' cases with heterogeneously end-competent voters .....	101
5.2.6	Unequal-stakes cases and the independence assumption .....	104
5.2.7	Unequal-stakes cases and direct causal dependence .....	108
5.2.8	Unequal-stakes cases and common cause dependence .....	117
5.3	Conclusions .....	122
6	Two more arguments and one alleged paradox .....	125
6.1	Introduction .....	125
6.2	The behavioural argument from weak collective optimality .....	125
6.3	The better-than-the-average-voter argument .....	128
6.4	Some remarks on the discursive paradox .....	131
6.5	Conclusions .....	136
7	Multi-option decisions .....	139
7.1	Introduction .....	139
7.2	The further-extended argument from collective optimality .....	142
7.3	Logrolling for self-interested voters .....	146
7.3.1	Is logrolling eventually against the logrollers' self-interest? .....	150
7.3.2	Can logrolling be avoided by devising compound decisions? .....	154
7.3.3	Why logrollers at best can achieve cycles or ties .....	156
7.3.4	Why logrolling in compound decisions is worse for logrollers and better for all (than in separate binary decisions) .....	162
7.3.5	Why logrollers in compound four-option decisions do not face the logrollers' dilemma .....	164
7.3.6	Results so far .....	167
7.4	Individual strategies for self-interested voters .....	168
7.4.1	Why any individual at best can achieve a cycle .....	170
7.4.2	Results so far .....	172
7.5	Alternative versions of the weighted majority rule .....	172
7.6	Common-interested voters and strategic voting .....	174
7.7	Less than fully competent voters .....	177
7.8	Conclusions .....	179
8	Summary and discussion .....	183
8.1	Introduction .....	183
8.2	The original argument can be adapted to other criteria of the common good .....	184
8.2.1	The original argument from collective optimality .....	184
8.2.2	The generic argument from collective optimality .....	185

8.3 The assumption of self-interested voting can be logically weakened .....	187
8.3.1 The extended argument from collective optimality .....	187
8.3.2 Arguments from weak collective optimality .....	189
8.3.3 The better-than-the-average-voter argument .....	194
8.3.4 The behavioural argument from weak collective optimality .....	194
8.4 The scope of the argument can be extended to multi-option decisions .....	195
8.4.1 The further-extended argument from collective optimality .....	196
8.5 Conclusion .....	199
Appendix .....	202
A.1 Introduction .....	202
A.2 Rae's argument for an individually optimal voting rule .....	203
A.3 How Rae's argument is relevant in the present context .....	205
A.3.1 The problem of bias .....	206
A.3.2 Deriving the weighted majority rule from the Raean framework .....	208
A.4 Two paradoxes of democracy .....	211
A.5 Conclusions .....	213
Bibliography .....	214



# Acknowledgements

This thesis would not have been written without the generous support from many sources. My greatest debt is to my two supervisors. I want to thank my main supervisor, Gustaf Arrhenius, for his acute and helpful comments on countless drafts of my thesis, for his good advice in philosophical and practical matters, and for his patience and support throughout the entire project. I also want to thank my second supervisor, Folke Tersman, for insightful and constructive comments on drafts of my thesis and on the project as a whole. Moreover, I am grateful to Krister Bykvist for his thorough and valuable criticism on a final draft of this thesis that turned out to be not so final after all.

I have had the good fortune to be accepted as a PhD-candidate to a department that provided me not only with a desk of my own, an income and administrative support, but also offers a lively and supportive philosophical community. I would like to thank the two consecutive heads of department, Björn Eriksson and Staffan Carlshamre, as well as Annika Diesen Amundin for their help and advice in many practical matters.

Special thanks goes to Niklas Olsson-Yaouzis for being such a great colleague, officemate and friend. Further thanks goes to all of my colleagues at the department, and especially Henrik Ahlenius, Hege Dypedokk Johnsen, Nicolas Espinoza, Lisa Furberg, Mats Ingelström, Sofia Jeppsson, Sandra Lindgren, Hans Mathlein, Jonas Olson, Anna Petrén, Daniel Ramöller, Maria Svedberg, Frans Svensson, Kjell Svensson, Torbjörn Tännsjö, and Olle Torpman. I would also like to thank my colleagues from the Department of Philosophy at Uppsala University, especially Per Algander, Emil Andersson, Karl Ekendal, Karin Enflo, Magnus Jedenheim-Edling, Jens Johansson, Victor Moberger and Karl Pettersson. I am grateful to all of them for philosophical advice as well as good company on numerous occasions.

In writing this thesis, I have benefited from a number of research visits to CERSES, Centre de Recherche Sens, Ethique et Société (Université Paris Descartes & CNRS), the Oxford Uehiro Centre for Practical Ethics (University of Oxford) and the Department of Philosophy at Humboldt-Universität zu Berlin. I am much obliged to Marc Fleurbaey, Roger Crisp, Julian Savulescu and Thomas Schmidt for making these visits possible.

Drafts of my thesis have been presented on numerous occasions: the PhD-seminars in practical philosophy at the Departments of Philosophy at Stockholm and Uppsala University, the colloquium for practical philosophy at Humboldt-Universität zu Berlin in 2012, the seminar in political theory at the Department of Political Science at Stockholm University in 2012, two workshops at the Swedish Collegium for Advanced Study (SCAS) in 2012 and 2011, two workshops of the Nordic Network on Political Theory in Oslo and Copenhagen in 2010 and 2009, and the Swedish National Philosophy Conference in 2009 and 2007. I am grateful to all the participants for their comments and constructive criticism. Special thanks goes to Katarina Barrling, Ludvig Beckman, Simon Birnbaum, Luc Bovens, Roger Crisp, Speranta Dumitru, Göran Duus Otterström, Jakob Elster, Eva Erman, Marc Fleurbaey, Max Fonseca, Gina Gustavsson, Frej Klem Thomsen, Kasper Lippert-Rasmussen, Christian List, Mats Lundström, Raino Malnes, Wlodek Rabinowicz, Bernard Reber and Thomas Schmidt.

I gratefully acknowledge the generous funding that made my research visits and workshop presentations possible. This was provided by the Royal Swedish Academy of Sciences, the Swedish Foundation for International Cooperation in Research and Higher Education (STINT), the Erik and Gurli Hultengren Foundation for Philosophy and the K & A Wallenberg Foundation. During the last months of this project, my work was funded by a generous stipend provided by the Helge Ax:son Johnson Foundation and a teaching opportunity at the Department of Philosophy at Uppsala University, kindly offered by Folke Tersman. For both I am very grateful.

Furthermore, I want to thank Michael Astroh, my first philosophy teacher and mentor, for opening my eyes to both the seriousness and the joy of doing philosophy.

My big German family deserves enormous thanks for their unfaltering encouragement and support in so many ways. I especially want to thank my parents, Sylvia and Ulrich Berndt, as well as Gisela Schulz, Susanne Schulz, and Christiane and Egbert Junghanns. I also want to record my gratitude towards my late grandparents, Christa and Joachim Berndt.

Finally, my warmest thanks goes to my children Felix and Hedvig, for tirelessly pointing out — both figuratively and literally speaking — what really matters, and to my fabulous wife Nika, for believing in me, supporting me, and for always being there, no matter what. I dedicate this book to her.

# 1 Introduction

## 1.1 Majority conundrums

In 2006, the City of Stockholm held a referendum on a heavily debated issue. The question was whether a congestion tax should be permanently implemented (after a trial period) for most of the inner city. There were two options: *yes* and *no*. Within the City of Stockholm, a clear majority of all valid votes (approximately 53%) were cast for *yes*.<sup>1</sup> The referendum was not legally binding, but merely advisory. Still, in 2007, the congestion tax was permanently implemented.<sup>2</sup>

However, many people in the surrounding municipalities within the County of Stockholm held the view that they should have been included. They claimed that the tax would affect them as well, as they commute or otherwise pass the city bounds regularly. As a result, 14 of the 25 municipalities held referenda, on their own initiative, on (roughly) the same issue. Considering all the County referenda in total, a majority of all valid votes was cast for *no* (approximately 60.2%). Putting together City and County votes would have resulted in a clear majority of valid *no*-votes (approximately 52.5%). So, arguably, there was a problem with the majority-based outcome of the City referendum: it was *due to a gerrymandered majority*, that is, a majority of an arbitrarily delimited group of people (City rather than County folks), or so one might claim. And this, it seems, makes the outcome arbitrary as well.

---

<sup>1</sup> All numbers are taken from the official website for *Stockholmsförsöket*:

<http://www.stockholmsforsoket.se/templates/page.aspx?id=10215> (accessed on 2012-06-25).

<sup>2</sup> The decision to permanently implement the congestion tax was taken by the Swedish government. Officially, the government based the decision not on the outcome of these referenda, but on other considerations. In order to justify that the government (rather than the City or County of Stockholm) owned the decision, it was stated that its outcome affected the entire nation (and not only City or County folks). And in order to motivate the decision to implement, it was claimed that this was in the interest of the nation. See the Swedish government's 2006 announcement 'Vi säger ja till trängselskatten för att finansiera kringfartsleder': <http://www.dn.se/debatt/vi-sager-ja-till-trangselskatten-for-att-finansiera-kringfartsleder> (accessed on 2012-06-25).

We are often faced with instances of majoritarian democratic decision-making.<sup>3</sup> In all these instances, the gerrymandered majority problem might arise. And there are other conundrums.<sup>4</sup> Consider the following scenario.

A local pub regularly hosts special nights for major sports events, showing televised competitions of all sorts on its big screen. Next Friday, there will be both an important golf tournament and a crucial football game, and the pub — having only one screen — can only show one of them. So, in order to accommodate the customers, the staff sets up a ballot box and asks them to vote on whether they want to watch the golf tournament or the football game. In addition, the staff asks all to vote on whether they would want to pay a small entrance fee for the event in exchange for cheaper prices on drinks and whether they would prefer smoking in the bar to be allowed on that night.

Now suppose that the pub's customers consist of three factions of roughly equal size: one faction desperately wants to see the football game, a second is strictly opposed to entrance fees and the third is allergic to smoking. However, on the respective issue, each faction is opposed by the other two factions, who slightly favour the other option. This means that each faction is outvoted on their most important issue — while getting their way on the other two, less important ones. So the pub will show the golf tournament, will take out an entrance fee in exchange for cheaper prices on drinks and will allow people to smoke. None of the three factions can stand it. So, in effect, no one will show up next Friday.

This is a version of the so-called *tyrannical majority* problem. On each issue, an almost indifferent majority of two thirds of the customers dominates a greatly affected minority. In such cases, the argument goes, a majority decision is problematic. This becomes especially clear when all three such decisions are considered: the majority gets its way on every issue, yet everyone opposes the combined outcomes.

Yet another problem emerges in the following example. Assume that, hypothetically, the French really had only three candidates to choose amongst in the 2012 presidential election: François Hollande, Nicolas Sarkozy and Marine Le Pen.<sup>5</sup> What would happen if slightly less than one third of the voters were left-wing voters, who rank Hollande over Sarkozy over Le Pen, slightly more than one third were right-wing voters, ranking Sarkozy over Le Pen over Hollande, while exactly one third were ‘protest’ voters, ranking Le Pen over Hollande over Sarkozy? If all voters voted for their top-ranked

---

<sup>3</sup> Of course, there are other rules in use, apart from majority rule. But in many decisions, at some point there will be an appeal to the will of the majority.

<sup>4</sup> Versions of the following two problems are also referred as ‘Democratic conundrums’ in Fleurbaey (*mimeo*: 29–32).

<sup>5</sup> A more complex example for all actual presidential candidates could be constructed, but would, in the present context, be unnecessarily difficult.

candidate in a direct vote, no candidate would gain majority support. Then there would be a second, run-off vote among the two candidates who got most votes in the first round. By assumption, this would be Sarkozy and Le Pen, since they are the first choice of the two slightly larger factions. If, among these two, every voter voted for the higher ranked candidate, Sarkozy would gain majority support (by a ‘coalition’ of left- and right-wing voters). However, it could then be reasonably complained that Sarkozy then would become president *against* the will of a majority of voters (namely, a ‘coalition’ of left-wing and protest voters) who rank Hollande above Sarkozy.

The problem can be brought out more clearly if we consider what would happen if majority rule was applied to each pair of presidential candidates. Then a ‘coalition’ of left-wing and protest voters would constitute a majority and thus select Hollande over Sarkozy. A coalition of left-wing and right-wing voters would constitute another majority and select Sarkozy over Le Pen, and finally, a coalition of right-wing and protest voters would form a third majority and select Le Pen over Hollande. Constructing a collective ranking from these majority outcomes would thus result in the following cycle: Hollande beats Sarkozy, who beats Le Pen, who in turn beats Hollande. This *cyclical majority* outcome implies that for each candidate, there is an alternative candidate who is supported by the majority. So majority rule would fail to select a winner.<sup>6</sup>

These examples show that majority rule at times runs into problems. Sometimes, it produces no outcome at all, as in the *cyclical majority* case. At other times, it produces a set of outcomes everyone opposes. This was the contention of the *tyrannical majority* case. Moreover, its outcome may be arbitrary, as when resulting from a *gerrymandered majority*.

Why are these problems for majority rule? As the cases have been stated, for every pair of options, people vote for whatever option they prefer, want, favour or rank higher. Now, suppose each does so because that particular option is in her self-interest, that is, *best for her* (among the two). Then, the option that most of the group — that is, a majority — vote for is the option that is best for most. This can be taken to mean that this option is *collectively best*, or in the common interest (among the two). Majority rule, since it selects this collectively best option as the collective outcome, can in this respect be concluded to be a good decision rule. And since it relies solely on an input in terms of the self-interest of the voters, it seems to be appealingly undemanding. The examples, however, point out several problems with this

---

<sup>6</sup> This case could also be spelled out as the *agenda problem* for eliminative voting rules such as pairwise majority rule where the loser of each round is eliminated among the remaining options. For the discussed hypothetical French presidential election, depending on which pair of candidates meets in the first round, the final outcome would change. This means that whoever controls the agenda would have more influence over the outcome than ordinary voters. In this study, I disregard agenda setting (cf. 3.4 below).

picture. It is hard to maintain that the outcome is collectively best if it can be shown to be opposed by all, arbitrary, or if in fact there is no definite outcome. Maybe majority rule is not so good then, after all.

It has been suggested that there is an improved version of the above considered *simple* majority rule that solves these problems: the *weighted* majority rule.<sup>7</sup> This rule operates as follows. It selects whatever option has received a majority of votes as winner, just as simple majority rule. But in contrast to the latter, the weighted majority rule assigns different amounts of votes to each person. More specifically, it assigns votes in proportion to what is at stake for each person in the given decision. This means that this novel rule rejects the classical motto of ‘one person, one vote’ (though it distributes votes equally in cases where everyone’s stakes are equal). Rather, the weighted majority rule assigns a large number of votes to people who have a lot at stake, a smaller number to those whose stakes are smaller, and no votes to those, and only those, who do not hold any stakes in the decision. If a person’s stakes are spelled out in terms of how much better one of the options is for her than the other — how much she is affected by the decision — the rule can solve the abovementioned problems in the following ways.

Consider first the referendum on the Stockholm congestion tax. The weighted majority rule would assign votes in proportion to every person who is affected by this decision. Arguably, it would thus assign votes to most City and Count folks (and some others as well). To the extent that all and only those affected should constitute the group of voters,<sup>8</sup> this means that by (correctly) employing the rule, an arbitrary delimitation of the group — gerrymandering — is avoided.

And consider the local pub sports event. As described, on each issue — whether to televise the golf tournament, whether to take an entrance fee and whether to allow smoking — there is a minority of voters with high stakes, who is outnumbered by a majority with small stakes. By (correctly) employing the weighted majority rule, the minority voters will, taken together, receive more votes than the majority voters, just in case the former’s stakes, taken together, outnumber the latter’s. In that case — which arguably is just the case of the ‘tyrannical’ majority — the minority’s votes outnumber the majority’s. Assuming that people vote for what is best for them, the outcome will then be best for each high-stake minority.

Finally, under the weighted majority rule (if properly employed), cycling majorities are rendered impossible. This means that the French — in the above slightly hypothetical presidential election — could expect a clear (or at least clearer) outcome. The argument showing immunity from cycling

---

<sup>7</sup> See e.g. Brighthouse and Fleurbaey (2010) and Fleurbaey (*mimeo*).

<sup>8</sup> One should reasonably add some other conditions: all and only those who are *relevantly* affected, *mature*, *capable* of voting and the like, should constitute the group of voters (that is, get votes).

requires a less sketchy account of the workings of the weighted majority rule. The argument will be stated in 7.2 below. (Even the arguments concerning the other two cases will be further clarified subsequently.)

It appears, then, that the weighted majority succeeds in deriving a common-interest outcome from purely self-interested input, without running into the problems faced by simple majority rule. The present study is an attempt to assess this claim. My overall goal is to analyse the workings of the weighted majority rule and to bring out the conditions under which this rule succeeds to select the collectively best, common-interest option as the outcome. (Note, though, that this is not a comprehensive study of the weighted majority rule: although I cover some grounds, I at times have to settle for simply pointing out loose ends and calling for further investigations.)

## 1.2 The ubiquity of the unequal vote

The weighted majority rule may initially seem to be an ill-suited solution to the described majority conundrums. Proposing an unequal vote may appear far-fetched, undemocratic or publicly unacceptable. However, a quick review of existing voting rules, which are usually considered democratic, reveals that these appearances are mistaken.<sup>9</sup>

First, to take a rather obvious example of real-life decision-making, the weighted majority rule is frequently applied in the context of shareholder democracy (also called corporate democracy). When shareholders are given the opportunity to vote on corporate policy, each shareholder's voting weights are proportional to the number of shares she holds in the company. This is a clear case of weighted voting according to stakes, when stakes are interpreted in terms of personal financial gain or loss. Within the corporate context, this does not seem undemocratic or unacceptable (even though it may be problematic from a larger societal or moral perspective).

Second, to take an example from the domain of politics, consider decision-making within the EU Council of Ministers. Each of the 27 European member states has one seat in the Council. Yet the voting weights for member states differ, being in (rough) proportion to their numbers of citizens. Thus the most populated states (Germany, France, UK, Italy) are currently assigned 29 votes each, while the least populated (Malta) has only three, with the other states ranging in between. It seems quite plausible that the underlying idea is that each member state representative votes in the interest of this state's citizens, and that the stakes are greater, the more citizens there are. In spite of occasional controversy on the specific voting weights, weighted voting is generally accepted and considered democratic in this context.

---

<sup>9</sup> Cf. Fleurbaey (*mimeo*: 33–37) for similar real-life examples.

Third, when people vote in real-life democratic elections, it is hardly ever the case that *every* person gets one vote, as the universal simple majority motto ‘one person, one vote’ demands.<sup>10</sup> Instead, equal amounts of votes are usually assigned to *confined groups* of people. This leaves many ‘outsiders’ with no vote at all. We tend to accept such unequal vote assignment whenever it mirrors whether people are relevantly affected by the decision in question or not. Thus, we tend to accept that the French are not given any votes in the Stockholm congestion tax referendum. On the other hand, we may complain that not all inhabitants of the surrounding municipalities within the County of Stockholm (or indeed all Swedes) were given a vote in this decision, if we believe all of them to be relevantly affected.

A corresponding motto of ‘one affected person, one vote’ can arguably be found in the widely accepted idea of subsidiarity. According to one intuitive interpretation of this idea, ‘decisions should be taken as closely as possible to the citizen’ in the sense that collective decisions should be made by the group of people (or their representatives) who ‘best approximates the set of relevantly affected people relative to the type of [decision]’.<sup>11</sup>

In these latter cases, however, vote assignment operates on an ‘on-off’ boundary between being affected by a decision and not being affected. It is insensitive to varying degrees of being affected. It may then be suggested that the ‘unequal’, weighted vote with such all-or-nothing weights in an important sense preserves equality, since it gives one vote each to the relevantly affected and zero votes each to those who are not relevantly affected. Thus it may be suggested that the weighted vote in such cases is not really an unequal vote and that this explains why it is used and generally accepted. (At least, one might add, this holds for contexts where stakes cannot be readily defined in terms of, e.g. financial gains or numbers of represented citizens, as in the above examples).

However, while it is true that weighted voting, which is sensitive to degrees of being relevantly affected, is rather unusual in real-life decision-making, it may not be unacceptable to many people. This is suggested by a recent experimental study.<sup>12</sup> The study shows that, under experimental conditions, people do tend to accept voting rules that are sensitive to varying degrees of being relevantly affected — and they do so to a greater degree than they tend to accept the rather insensitive simple majority rule. In the experiment, participants were presented with a hypothetical case of city residents who could vote for or against a city construction site (for an industrial or housing complex). The city residents were described as holding different stakes on this issue, expressed either in terms of how much their apartments would increase or decrease in value, as a result of the construction, or in

---

<sup>10</sup> Cf. Fleurbaey (*mimeo*: 32–33).

<sup>11</sup> Arrhenius (2013: 7).

<sup>12</sup> Dimdins et al. (2011).

terms of how close to the construction site they lived. The participants of the experiment were then asked to rank different voting schemes for this case, in light of their own ‘personal views of fairness, wisdom, and the greater good of society’.<sup>13</sup> The voting schemes included simple majority rule (assigning one vote to each city resident) and different weighted majority rules (assigning numbers of votes, in different proportions to the residents' stakes). The analyses of these rankings ‘clearly show that participants preferred voting schemes that positively differentiated between groups with different stakes, assigning more voting power to groups with higher stakes’.<sup>14</sup> There are also some results suggesting that more information about residents' stakes increased the participants' acceptance of weighted majority rules. (These results, however, rely on the fact that the participants accept the definition and assessment of the residents' stakes, as the authors point out.<sup>15</sup>)

So it seems that weighted voting is neither a far-fetched idea, nor publicly unacceptable or undemocratic from the outset. It may be interesting and worthwhile to investigate weighted voting rules more closely. The present study attempts to do just that for a specific voting rule, which here is called the weighted majority rule.<sup>16</sup>

### 1.3 The value of democracy

It should be noted that the framing of the above majority conundrums and their proposed solution rests on a substantial idea concerning the value of democratic decision-making or democracy. (I here use these terms interchangeably.) The general idea is that such decision-making is valuable in so far as it selects the collectively best option, that is, the option that is in the common interest. Democracy is thus taken to be of instrumental value. Moreover, the common-interest option was claimed to be the option that is ‘best for most’. If the notion of ‘being best-for someone’ is interpreted in terms of ‘comprising most individual well-being’ (in some sense), well-being is understood as a constitutive part of the common interest. So the idea, one could claim, is that democracy is instrumentally valuable since it maximises individual well-being, aggregated across the entire group.

There are a number of other ideas about the value of democracy. Some political philosophers argue that it embodies or realises *fairness* or social *equality*, since it treats everyone with equal respect or grants them equal

---

<sup>13</sup> Dimdins et al. (2011: 19).

<sup>14</sup> Dimdins et al. (2011: 7).

<sup>15</sup> Dimdins et al. (2011: 9).

<sup>16</sup> There is a recent upsurge in interest, within democratic theory, in a variety of voting rules with an unequal or weighted vote. I refer to some papers below, in 2.2.1.

political power.<sup>17</sup> Others claim that it realises *liberty* or personal *autonomy* among the participants, since it allows them to decide for themselves.<sup>18</sup> According to these ideas, democratic decision-making is inherently valuable: it derives its value from the (alleged) value of the things it realises in a non-causal way. In contrast, others argue that democracy has good causal consequences. They see it as instrumentally valuable. According to some of them, it promotes a more *virtuous character* among its participants.<sup>19</sup> According to others, it produces *outcomes that respect alleged values*, such as liberty or equality.<sup>20</sup> And still others bring forth arguments, similar to the present study, that democracy yields *better outcomes in terms of well-being*. These views are neither mutually exclusive nor is their list exhaustive. They merely illustrate different ways of understanding the general idea that democratic decision-making is good.

The claim that democracy maximises collectively aggregated well-being (in some sense) can be supported by different kinds of arguments. One can set out to *empirically* measure individual levels of well-being and assess how their aggregate correlates with the degree of democracy (in some sense) within the respective context. Such studies have been claimed to show, for instance, that democracies (that is, states in which democratic decision-making has some important role) to a lesser degree than other forms of states experience wars or famines, or that their citizens tend to have higher well-being.<sup>21</sup> These results may then be followed up by an analysis of why this is so. It might be a psychological fact that people are happier or more satisfied when they are given influence over collective decisions. Or maybe in democratic states people face other, better options to choose amongst, e.g. because

---

<sup>17</sup> For an early account of democracy in terms of equality, see Tocqueville (1990: 9, 19). For a contemporary account of democracy as ‘a paradigm of a fair compromise’, see Singer (1973: 32). For an account of democracy as the unique (public) realisation of equality, see Christiano (1996: chapter 2). Cf. also Dahl (1989).

<sup>18</sup> An early defender of democracy in terms of freedom or autonomy is Democritus who claims that ‘poverty under democracy is as much to be preferred to so-called prosperity under an autocracy, as freedom is to slavery’ (quoted in Naess et al. 1956: 79). For a contemporary (critical) account of self-government as a foundation of democracy, see Christiano (1996: chapter 1). For defences, see e.g. Gould (1988: 45–85). For the view that democratic processes ‘express the autonomy and equal standing of citizens’, see Anderson (2009: 225).

<sup>19</sup> For character improvement, see e.g. Mill (Considerations on Representative Government: 74). See also Elster (2002: 152).

<sup>20</sup> For a sketch of an instrumentalist defence of democracy in terms of equality, see e.g. Arneson (2009). See even Dworkin (1987). On the promotion of liberty, see e.g. Nelson (1980).

<sup>21</sup> See e.g. (Sen 1999a: 152): ‘no substantial famine has ever occurred in any independent country with a democratic form of government and a relatively free press’. For recent criticism, see Rubin (2009). See Henderson (2002: 3) for a bibliographical review and critical treatment of the ‘democratic peace proposition [...] that democratic states are less likely than nondemocratic states to fight wars against each other’. See e.g. Przeworski et al. (2000) for an empirical assessment of the impact of democracy or dictatorship on well-being.

they have influence over the agenda. Or again, among the given options, people manage more often to select the best outcome when they vote democratically.

Another kind of argument for the claim that democracy promotes some value, such as aggregated well-being, focuses not so much on empirical data, but rather on *formal models* of democracy. These models can be described as consisting of three main components: a democratic decision rule, assumptions about the input for this rule and conclusions about the output of the rule (given the input). In other words, the rule is considered as a mechanism that derives an output — a collective choice or ranking — from an input — such as individual votes, voting weights, preference rankings and the like.<sup>22</sup> The point of doing this is so that one can isolate certain features within the rule, the input and the output, and study how they relate to each other. Formal models are thus understood as an analytical tool that allows one to see details in the workings of a democratic decision rule, which are usually buried in the complexity of real-life decision-making.

In order to build an argument for a specific democratic decision rule from such a model, one can, from specific assumptions about the input — individual votes or the like — derive the output — a collective choice or ranking. One can then evaluate the latter in the light of a normative criterion, such as the criterion of maximising collectively aggregated well-being. Two versions of such a formal model argument have been recently proposed for the weighted majority rule, by Marc Fleurbaey and by Harry Brighouse and Fleurbaey, respectively.<sup>23</sup> The argument shows that this rule, when there are two given options and voters vote in their self-interest, selects the collectively best option, according to two interpretations of this normative criterion. I call this the *original argument from collective optimality*. My study is an attempt to reconstruct, critically examine and improve this argument, mainly focusing on the conditions, or assumptions, on which it rests.

## 1.4 Main thesis and disposition

The main thesis of this study is that the *original argument from collective optimality* can be improved, in the sense that its assumptions can be logically weakened. That is, I claim that the assumptions of the *original argument* imply — but are themselves not implied by — the assumptions of my improved arguments. This means that, within the set of all possible cases of collective decision-making, the *original argument* is relevant for a set of cases that is a proper subset of the set of cases for which my improved ar-

---

<sup>22</sup> Different versions of these models are employed by e.g. Arrow (1963), Downs (1957), Black (1958) and Buchanan and Tullock (2004). More on those in 2.5.2 below.

<sup>23</sup> Fleurbaey (*mimeo*) and Brighouse and Fleurbaey (2010).

guments are relevant. In other words, the collective optimality of the weighted majority rule can be established for a larger set of possible cases of collective decision-making than previously thought.<sup>24</sup>

My arguments for the main thesis proceed along three main lines of inquiry. The first concerns the question whether the *original argument* can be adapted to other normative criteria than the ones suggested by Brighouse and Fleurbaey. In Chapter 2 I present the basic concepts and theories underlying this study. I define the weighted majority rule and state alternative theories of self-interest (and individual well-being) and of the common good (and the common interest). I moreover relate my study to the relevant literature and comment on its methodology.

In Chapter 3 I reconstruct the *original argument from collective optimality*, stating the assumptions under which the weighted majority rule is collectively optimal, in the sense of selecting the common-interest option. I show that the argument can be adapted to different criteria of the common good and propose a *generic* version of the argument that works for a number of these criteria. I moreover state some limits and clarifications of the argument.

My second line of inquiry constitutes the main part of this study. I pursue the question of whether the assumption of self-interested voting can be logically weakened while preserving the collective optimality of the weighted majority rule. In Chapter 4 I show that the original argument can be *extended* to allow even common-interested voting. I then consider this relaxed assumption of self- or common-interested voting in greater detail and suggest that it is implied by a set of assumptions concerning the voters' motivating desires (to pursue their self-interest or the common interest) and relevant beliefs about the options (in the light of these desires).

In Chapter 5 I address the resulting assumption that voters have *correct* beliefs about their self-interest or the common interest. Employing a number of so-called Condorcet jury theorems, I show that this assumption can be considerably relaxed under certain additional conditions, such as that there are large numbers of voters. In such cases, the weighted majority rule can be shown to be collectively optimal even if voters are (on average) only somewhat better than chance at correctly judging the options. However, collective optimality must here be understood in a weaker sense, as 'selecting the common-interest option with certainty or near certainty'. This chapter thus states *arguments from weak collective optimality*.

In Chapter 6 I use the results of the previous chapter to devise two additional arguments. The *behavioural argument from weak collective optimality*

---

<sup>24</sup> Note that it is a further question whether this means that the improved arguments are relevant for a larger set of *actual* cases of collective decision-making, compared to the original argument. It might, after all, be the case that of all the possible cases for which they are relevant, none will ever be realised.

relies on an assumption concerning voting behaviour, regardless of the motivational set-up underlying or implying such behaviour. The *better-than-the-average-voter argument* shows that even for smaller groups of voters, the weighted majority rule still has some attractive features. A brief discussion of the so-called discursive paradox brings out how framing the options can affect the voters' probabilities of correct beliefs.

My third line of inquiry concerns the question of whether the scope of the *generic* and *extended arguments from collective optimality*, as discussed in Chapters 3 and 4, can be *further-extended* to decisions with more than two options. In Chapter 7 I construct an argument showing that this is the case. I then point out that the voting behaviour assumption on which this argument rests is ambiguous in multi-option settings. I argue that once disambiguated, the *further-extended argument from collective optimality* cannot be stated on the more plausible interpretation of the assumption, since on this interpretation, strategic voting becomes possible and undermines the collective optimality of the weighted majority rule. I analyse two forms of strategic voting, practiced by groups of voters ('logrolling') and by individual voters ('individual strategies') respectively. I examine to what extent these forms of strategic voting undermine the collective optimality of the weighted majority rule and to what extent they benefit self-interested and common-interested voters who are either certain to judge the options correctly or at least better than chance. A discussion of alternative versions of the weighted majority rule brings out the need for further research to refine this rule.

Chapter 8 summarises the main results of this study, along the three stated lines of inquiry. I discuss some of the implications and speculate about the practical relevance of my results and illustrate them on occasion with some of the 'majority conundrum' cases described in this introduction.

Finally, in the Appendix I consider a question that is closely related to my main thesis yet does not constitute an argument for it. The question is whether a self-interested voter could accept the weighted majority rule as a method of collective decision-making. Starting out from a well-known argument to the effect that self-interested voters have reason to accept the *simple* majority rule, I show that when the assumptions of this argument are improved, it advocates the *weighted* majority rule instead.



## 2 The basics

### 2.1 Introduction

The main thesis of this study concerns an argument from collective optimality, interpreted in terms of collectively aggregated well-being, for the weighted majority rule. In section 2.2 of this chapter, I spell out and illustrate this rule in greater detail. Moreover, in 2.3 I spell out what is meant by ‘collective optimality’ and the related concept of ‘individual well-being’. Section 2.4 gives some theoretical background regarding the relationship between democracy and the common good. Finally, section 2.5 sketches the philosophical background of the entire project and comments on its methodology.

### 2.2 The weighted majority rule

The weighted majority rule is a rather unorthodox democratic decision method, recently proposed in two separate papers by Brighthouse and Fleurbaey and, separately, by Fleurbaey.<sup>1</sup> In the standard case, the rule is applied to decisions with two options. I call these *binary* decisions. The weighted majority rule states that every person's vote is to be assigned a voting weight in proportion to what is at stake for this person in the decision and that the option that receives more voting weights is selected as the outcome, or winner, of the collective decision.

There are a number of possible ways to spell out ‘stakes’ here. One idea is to define it simply as the difference in well-being between the two options for the person in question.<sup>2</sup> Then, if one of the options makes you quite well off, while the other brings you down to an extremely low level of well-being, your stakes are quite high and thus your assigned voting weights rather heavy. (Of course, someone else might have even more at stake and hence even heavier voting weights.) If, on the other hand, both options make you

---

<sup>1</sup> Brighthouse and Fleurbaey (2010) and Fleurbaey (*mimeo*).

<sup>2</sup> Cf. Fleurbaey (*mimeo*).

equally well (or badly) off, you have no stakes in the decision and hence receive no voting weights.

Another idea is that ‘stakes’ are defined as the difference between the *weighted* levels of well-being for the two options, with the weights chosen in proportion to how badly off the person in question is made by either option.<sup>3</sup> Then, you and I could have the same difference in well-being between the two options, but you would still receive heavier voting weights than I would if you were overall worse off than me (that is, if your well-being differential were located on a ‘lower end’ of the welfare scale than mine).

There are many other ways to spell out ‘stakes’. I return to this issue in 3.2 and 3.3 below. Meanwhile, for the sake of simplicity, I stick to the first suggestion, defining a person's stake simply as the difference in well-being between the two options for the person in question.

Before turning to an example of how the weighted majority rule is applied, one further note is necessary: as stated, the weighted majority rule assigns varying voting *weights* to votes. I find it easier to present my examples and arguments in terms of varying *numbers* of votes, so this is what I do in the remainder of this study. But this way of speaking should not suggest that voters could split their votes between different options. Rather, when I speak of numbers of votes assigned to some voter, this should be understood as an *indivisible vote bundle*.

***The Weighted Majority Rule:*** For all individuals and any decision with two options, (a) every individual is assigned a number of votes in proportion to her stakes, and (b) the option that receives a majority of votes is selected as outcome.

The following pair of scenarios provides a simple illustration of the workings of the weighted majority rule, thus understood.

***Dinner plans 1.*** A group of three, Abby, Beth and Charlie, needs to decide whether to eat out at a restaurant or stay in and cook for themselves tonight. They therefore have two options: *out* and *in*. Abby is made better off with *out*, as is Beth, while Charlie is better off with *in*. Let us assume that they all have the same stakes, that is, that they have an equal amount of well-being to gain. This means that the weighted majority rule will assign the same number of votes to all. Abby casts her, say, one vote for *out*, as does Beth, while Charlie casts her one vote for *in*. Thus *out* receives the majority of votes and is selected. The three will eat out tonight.

In this equal-stakes scenario, the *weighted* majority rule operates exactly as the *simple* majority rule. It assigns one vote to each voter. Now, imagine a slightly different scenario.

---

<sup>3</sup> Cf. Brighouse and Fleurbaey (2010).

**Dinner plans 2.** Within the above group of three, Charlie's stakes are three times as high as Abby's or Beth's. What this means is that there is three times as much well-being at stake in the decision for Charlie as is for Abby or Beth. Let us simply suppose that, while they are all just about equally hungry, Charlie is nearly broke, and eating out — contrary to eating in — is too expensive for her to buy enough to eat. The more affluent Abby and Beth will be sure to still their hunger irrespective of whether they eat in or out; they will be just slightly better off from eating out. Then Charlie, who votes for *in*, will by the weighted majority rule receive, say, three votes, while Abby and Beth, voting for *out*, will receive only one vote each. Now, *in* will get the majority of votes and thus be selected instead of *out*. The three will eat in tonight.

So in this unequal-stakes scenario, the weighted majority rule differs from the simple majority rule. One voter gets more votes than the others, simply because there is more at stake for her.<sup>4</sup> As a result, the outcome changes.

Some might question whether the weighted majority rule is a democratic decision rule. I address this worry in the next section.

### 2.2.1 The weighted majority rule and democratic theory

The weighted majority rule does not seem to qualify as a democratic voting rule according to two very common ways of defining the latter. One common definition focuses on equality and is often referred to under the already quoted motto ‘one person, one vote’.<sup>5</sup> The other defines democracy in terms of majority rule, in the sense of requiring that (at least) a majority of the voters prevail over the minority.<sup>6</sup>

---

<sup>4</sup> Note that in the example all vote according to their stakes, yet they may have voted otherwise. The weighted majority rule does not commit voters to self-interested voting. However, as we will see in Chapter 3, the *original argument from collective optimality* for the weighted majority rule makes an assumption to this effect.

<sup>5</sup> Cf. e.g. Christiano's (2008: §1) definition: ‘To fix ideas, the term “democracy”, as I will use it [...], refers very generally to a method of group decision making characterized by a kind of equality among the participants at an essential stage of the collective decision making’. Cf. even Barry's (1991: 25) definition: ‘By a democratic procedure I mean a method of determining the content of laws (and other legally binding decisions) such as that the preferences of the citizens have some formal connection with the outcome in which each counts equally’. Note though that the requirements of ‘a kind of equality among the participants’ or of ‘each count[ing] equally’ may be argued to be satisfied when we give equal consideration to the individuals' *stakes* — rather than to individuals regardless of their stakes. See my next paragraph.

<sup>6</sup> Tännsjö (1992: vii), for instance, defines democratic decision-making in terms of the will of the people, and the latter in turn in terms of majority will, and states: ‘[...] I define the concept of democracy in classical terms, in terms of the majority principle and the principle of unanimity [...]’. Hardin (1993: 158) writes: ‘In modern political thought, the core of the notion of democracy is its etymological core — rule by the people — which translates most naturally as majority rule if there are divisions of opinion’. Dahl (1989: 135) states that ‘virtually everyone assumes that democracy requires majority rule in the weak sense that support by a majority ought to be necessary to passing a law. But ordinarily supporters of majority rule mean it in

The weighted majority rule, as we have just seen, neither gives every voter an equal vote in the decision nor lets the majority of voters prevail. Still, I wish to present it as a democratic voting rule. One reason is that the rule does fulfil the equality and majority requirements in an important sense — which pertains not to voters, but to stakes. Its basic features can be re-described as, first, assigning to every stake an equal voting weight and, second, letting the majority of votes — as distributed in proportion to stakes — prevail.<sup>7</sup>

A second reason is that there are a host of alternative definitions of democratic decision rules that arguably could vindicate the weighted majority rule as democratic. This is true, e.g. of the class of definitions that define democratic rules in terms of popular control, along the lines of the ‘Lincoln formula’ for government ‘of the people, by the people, for the people’.<sup>8</sup> The weighted majority rule is a collective decision rule (governing ‘the people’), which derives a common-interest outcome (‘for the people’) from individual input by vote (‘by the people’). It thus seems to match the Lincoln formula.

Third, as William Riker remarks, participation by voting is the common element of most definitions of democracy as well as ‘the central act of democracy’.<sup>9</sup> Similarly, David Estlund claims that the ‘core’ of the idea of democracy is to ‘rule by the people by way of voting’.<sup>10</sup> The weighted majority rule stays true to this essential commitment to the popular vote.

It should moreover be noted that the idea of an unequal vote is not new to democratic theory. There is, for instance, a long tradition of theorising about the assignment of voting weights in proportion to competence. This idea is often associated with John Stuart Mill, who proposed voting weights in proportion to the degree of education or intelligence.<sup>11</sup> It has its modern defenders as well. On modern accounts of the competence-weighted majority rule, it states that every voter's vote is weighted by the logarithm of her probability of voting correctly.<sup>12</sup> This rule is proposed primarily within the so-called

---

[the] much stronger sense [...] that majority support ought to be not only necessary but also sufficient for enacting laws.’ Note though that the requirement of ‘majority’ support may be argued to be satisfied when a majority of *votes* — distributed to voters in proportion to their stakes — prevails, rather than a majority of voters. See my next paragraph.

<sup>7</sup> A further reason is that, if we start from the unquestionably democratic (according to both the equality and the majority requirement) simple majority rule and try to improve it from an individual optimality perspective, we arrive at the weighted majority rule (or so I argue in the Appendix).

<sup>8</sup> Naess et al. (1956: 37). As the authors point out, Abraham Lincoln in his Gettysburg Address did not use the phrase as a definition of ‘democracy’. As e.g. Harrison (2005: sec. 1) states: ‘Democracy means rule by the people. [...] The people themselves rule and they rule themselves. The same body is both ruler and ruled’.

<sup>9</sup> Riker (1982: 5).

<sup>10</sup> Estlund (1990: 397).

<sup>11</sup> Mill (1861: chapter VIII).

<sup>12</sup> See e.g. Nitzan and Paroush (1982) and Shapley and Grofman (1984).

‘judgment-aggregation approach’ to democracy (which I outline briefly within this section, see below).

There are also debates on assignment of voting weights in proportion to other things, such as the number of owned shares in a company within corporate democracy (we may call this the *share*-weighted majority rule),<sup>13</sup> or the number of represented constituents within indirect democracy (we may call this the *representation*-weighted majority rule).<sup>14</sup> In this study, I focus on the *stake*-weighted majority rule and disregard alternative grounds for weight assignment. In the following text, ‘weighted majority rule’ is a shortcut for ‘stake-weighted majority rule’, unless stated otherwise.

There has also been some recent theoretical interest in alternative voting rules that, generally speaking, take the voters' varying stakes (or ‘preference intensities’) into account in other ways, and are, because of this feature, welfare-promoting. For instance, Rafael Hortala-Vallve analyses a so-called qualitative voting rule. This rule assigns an equal number of votes to every voter to be distributed freely among several binary issues. By using more votes on certain issues, voters can express greater stakes on these, compared to other issues.<sup>15</sup> Alessandra Casella studies a storable vote rule. The rule endows each voter with one vote per decision, letting her either use it directly or store it for future decisions. Storing votes from lower-stake decisions allows the voter over time to concentrate them on decisions in which her stakes are higher.<sup>16</sup> David Heyd and Uzi Segal propose a two-tier voting rule. In the first stage of this rule, everyone is asked to assign a voting weight to each prospective voter, according to certain considerations. In the second stage, everyone is then asked to vote on the binary issue in question, with each voter's vote being counted in accordance with her average weight, as it results from the first stage. The first-stage weighing, the authors propose, is done in the light of considerations ‘based on the interests of the people who are going to be affected by the policy’, or ‘based on the cognitive position of the voters making the decision on that policy’.<sup>17</sup> The rule can thus take into

---

<sup>13</sup> See e.g. Leech (2001). Leech points out the failure of weighted voting to allocate voting power in proportion to number of owned shares. I am, however, not concerned with voting power. Cf. even Brighouse and Fleurbaey (2010: 145).

<sup>14</sup> See e.g. Banzhaf (1965) and Felsenthal and Machover (1997). They discuss the failure of weighted voting to allocate voting power in proportion to the number of represented constituents. See also Barberà and Jackson (2006). Cf. Beisbart and Bovens (2007: 582f.), who argue that ‘representatives of interest groups [that is, groups in which the interests of the people fully overlap] should have [...] weights proportional to the sizes of their respective interest groups on the utilitarian ideal’. This comes quite close to the idea that voters (as ‘representatives’ of their stakes) should have stake-proportional voting weights given e.g. a sum-total criterion of the common good (cf. 3.2 below).

<sup>15</sup> Hortala-Vallve (2012).

<sup>16</sup> Casella (2005).

<sup>17</sup> Heyd and Segal (2006: 107)

account (estimates of) stakes as well as competence in assigning voting weights.

These rules may prove to be more or less collectively optimal, in my sense of the term. However, I focus my study on the (stake-) weighted majority rule.

I now briefly outline how my discussion of this rule fits into, and distinguishes itself within, a larger framework of democratic theories. The weighted majority rule is here considered within what may be called a populist, preference-aggregative account of democracy. There are other, alternative approaches. I briefly outline three important candidates, state how they differ from this account and indicate how some of the differences are reconciled in this study.

***The deliberative approach.*** This approach makes democracy not so much a question of the individual act of voting, but rather one of the public, joint practice of *deliberation*. The basic idea is that the members of a group publicly propose and debate their conceptions of the common good, as well as publicly acceptable reasons for these conceptions, to persuade each other as rational equals. The contention is that in general discussion participants have a desire to reach agreement and are constrained by the condition of publicity, and thus cannot appeal to idiosyncratic views but must rather argue from potentially generally acceptable principles. This means that the democratic input is transformed and arguably improved by the constraints of rationality and publicity. Ideally, deliberation results in unanimous agreement on a common conception of the common good. Non-ideally, voting is the last resort to resolve remaining conflict. Whatever the outcome, the idea is that it will in some sense be superior to the mere aggregation of unqualified individual input.<sup>18</sup>

In contrast to the deliberative approach, my account of democracy has *aggregation* as its central feature. These two are not mutually exclusive though. For most accounts of democracy, the question of aggregation versus deliberation is rather a question of degree or priority. While most deliberative accounts in the end also rely on vote aggregation, most reasonable aggregative accounts will reserve an important theoretical place for deliberation and improvement of the individual input. (I briefly return to this issue in 5.2.2 below.)

***The judgment-aggregation approach.*** This approach does have aggregation as one of its central features. More exactly, it claims that individual judgments are aggregated in a way that ‘tracks the truth’. How does this work?

---

<sup>18</sup> See e.g. Cohen (2003). Cf. even Anderson (2009: 217), Elster (2003), Dryzek (2000) and Gutmann and Thompson (1996). For a critical account of deliberative democracy (or ‘the constructive view’), see Christiano (1996: 35–43).

According to Joshua Cohen, it is first assumed that there is ‘an independent standard of correct decisions — that is, an account of [...] the common good that is independent of current consensus and the outcome of votes’. Moreover, it is assumed that votes express judgments ‘about what the correct policies are according to the independent standard, not personal preferences for policies’. These judgments of the options thus either do or do not conform to the independent standard. Then, as is shown by the influential Condorcet jury theorem, under certain conditions there is a good chance that the outcomes of democratic decision-making are correct judgments of the options, that is, judgments that conform to the independent standard. (I introduce the theorem properly in 5.2.1 below.) This means, then, that the approach proposes ‘an account of decision making as a process of the adjustment of beliefs, adjustments that are undertaken in part in light of the evidence about the correct answer that is provided by the beliefs of others’.<sup>19</sup>

It is common to distinguish this *judgment*-aggregation approach from *preference*-aggregation accounts similar to the one I am concerned with.<sup>20</sup> The first assumes votes to express judgments about the given options in the light of the voter's perception of a *common* standard, such as truth or objective correctness. In contrast, the other assumes a vote to express the voter's perceived preferences concerning the options. That is, the voter ranks the options e.g. according to her perception of what she likes best, or what is best for her. Since the ranking criteria might differ between different voters, we can call this a voter's *individual* (rather than a common) standard. The difference between these two approaches is then that, had all the voters a correct perception of whatever they vote on, they would all vote for the same option on the first approach, but possibly for different options on the second.

Of course, the idea that the votes reflect different individual standards can be combined with the idea that there is some common standard as well, in the light of which these votes can be evaluated. I get back to a similar idea below (in Chapter 5) and explore how the judgment-aggregation approach and its analytical tools (different versions of the Condorcet jury theorem) can enhance our understanding of my present account of democracy.

***The elitist approach.*** This approach is best understood as a reaction against classical accounts of democracy that seek to justify democracy because it serves the common good. In Joseph Schumpeter's words, the complaint is that the common good is a chimaera, due to people's ‘irreducible differences of ultimate values which compromise could only maim and degrade’, as well as people's disagreement about the proper means to any end that can be accepted by all. Moreover, Schumpeter argues that the aggregation of the individual input must be arbitrary since there is no independently justified meth-

---

<sup>19</sup> Cohen (1986: 34), italics omitted.

<sup>20</sup> See e.g. Rabinowicz (*mimeo*).

od of derivation and since the individual inputs are plagued by irrationality, mistakes and manipulability. In the face of these difficulties, he insists that the individual voters' role be kept to a minimum. Their primary function is to periodically elect — and by extension remove — governors. Thus, 'the democratic method is that institutional arrangement for arriving at political decisions in which individuals acquire the power to decide by means of a competitive struggle for the people's vote'.<sup>21</sup>

This means that in the elitist approach, voters vote on an extremely limited number of issues. They merely decide who should in effect decide all substantive issues. In contrast, my present account of democracy can be called populist, since it gives voters a key role in virtually all decisions that concern them. Specifically, the weighted majority rule is not committed to limiting the domain of issues to the election of leaders but is in principle applicable to any issue — from what to eat for dinner to what we should do about global warming. Nor does it limit the demos to citizens of nations (or the like) but considers any collective as eligible for demos — from small groups (say, a group of three, or the crowd that just now happens to be gathered in the local pub) to large-scale assemblies (such as the entire nation of France, or all human beings).

The scope of the present account, both concerning potential issues and potential demoi, is thus extended beyond the traditional domain of politics, in most of its definitions. This is as it should be, since the aim of this study is to analyse a voting rule from the perspective of the common interest, which obviously is not constrained by any such domain.<sup>22</sup>

Let us go back to the weighted majority rule within the outlined account of democracy. In order to understand it better, we still need to get a better grasp of how 'stakes' should be understood. The weighted majority rule, it was stated, assigns to each voter a number of votes in proportion to her stakes. Stakes are in the present study linked to individual well-being. I now want to address the question what individual well-being is. The question of exactly how the stakes are defined is answered later (see 3.2 and 3.3. below).

---

<sup>21</sup> Schumpeter (1975: 251–254; 269). For the main objections to the Schumpeterian account, see Christiano (1996: 134–140). For a reading of Schumpeter's account in descriptive terms, see Arrhenius (2013). Schumpeter's account is closely related to Riker's 'liberal view' that 'the notion that voting permits the rejection of candidates or officials who have offended so many voters that they cannot win an election' (Riker 1982: 242). For a rigorous criticism of Riker's conclusions, see Mackie (2003: *passim*).

<sup>22</sup> The distinction between the political and the personal, between public and private is as untenable from such a general perspective of the common interest, as it is from a feminist perspective. See e.g. MacKinnon (1989: 191).

## 2.3 Individual well-being and the common good

The main thesis of this study concerns an argument from *collective optimality*. The notion of ‘being collectively best’ was somewhat sketchily introduced as ‘comprising most collectively aggregated well-being’. I now briefly describe some important philosophical theories of individual well-being, as well as theories concerning how individual well-being relates to the common good, or common interest. The former are theories about what is good for an individual, what makes her well off or is in her self-interest. The latter concern, in a way of speaking, what is better or worse for a group of individuals, what is the common good or in the common interest.

I do not take a stand on which of these theories is the correct or most plausible one. This task is beyond the scope of this study, which focuses on the weighted majority rule. More importantly, I do not want to limit the relevance of my study unnecessarily. Instead I want to employ ‘well-being’ and ‘common good’ as placeholders for whatever theories the reader has in mind — within certain constraints, as stated at the end of section 2.3.2.

### 2.3.1 Individual well-being

Philosophical theories of well-being are usually classified in three main approaches: hedonistic, desire-fulfilment and objective list theories.<sup>23</sup>

**Hedonism.** According to classical hedonism, well-being is happiness, which in turn consists in the balance of pleasure over pain.<sup>24</sup> A formal or explanatory version of hedonism says instead that what makes something good for someone is pleasantness, what makes something bad for someone is painfulness. Both state that the more intense, or the longer in duration the pleasantness (painfulness), the greater (lesser) the well-being.

According to some hedonist theories, pleasantness and painfulness are mental states, more precisely, sensations: they are the positive or negative ‘feeling tone’ shared by all experiences that we find pleasant or painful.<sup>25</sup> Others maintain that they are desired and undesired consciousness respectively. Thus they can consist of a multitude of different experiences that only share the common denominator of being the objects of certain attitudes of the agent. According to these theories then, what makes something good (bad) for someone is the agent’s attitude toward her experience of it.<sup>26</sup>

All hedonist accounts of well-being subscribe to the Experience Requirement: that well-being consists only in the relevant mental state (pleasantness

---

<sup>23</sup> This goes back to Parfit (1984: 3).

<sup>24</sup> This is Bentham’s (2000) position.

<sup>25</sup> For critical discussions, see Crisp (2006), Griffin (1988: chapter 1), Sumner (1999: chapter 4) and Tännsjö (1998: chapter 5).

<sup>26</sup> For critical discussions, see Brandt (1998), Sumner (1999) and Feldman (2004).

or the relevant attitude). They are thereby vulnerable to Robert Nozick's infamous Experience Machine objection.<sup>27</sup> If mere experience of the relevant mental state is all that is required for making some state of the world good for us, should our plugging into a simulator, which simply gave us these experiences in form of neural stimulations, not make our lives so much better? Yet this seems utterly implausible — or so the objection goes. This is then taken to show that there is something wrong with all hedonist accounts of well-being. However, the soundness of the Experience Machine objection has been heavily debated.<sup>28</sup>

***Desire-fulfilment theories.*** These theories attempt to get out of the mind and into the world, but without losing touch with people's own priorities. Desires seem like a good place to start — desires not just for experiences themselves, but for the external states experiences (usually) represent. Still, having one's desire fulfilled does not necessarily make one better off. I may desire to drink a glass of clear liquid, being unaware of the fact that it contains poison, not water. Or I may desire being famous, yet be unaffected by the desired state of the world, e.g. if it is realised after my own death, after I have lost the desire or without my knowledge. Most desire theories seek to restrict the set of desires whose satisfaction can make us better off, requiring e.g. that desires be informed, be the result of cognitive psychotherapy, be held at the time of their satisfaction, be about the desirer's own life or that their satisfaction be experienced by the desirer.<sup>29</sup> Such restricted desire theories then claim that what makes life good for someone is qualified (in the relevant senses) desire-fulfilment.

One main objection is that there is something very strange with the claim that desire fulfilment (even of the qualified sort) is a good-making property. This would mean that things are good because we desire them, while it seems far more plausible to claim that we desire things because they are good. The question of whether desire accounts implausibly reverse the order of explanation from value to desire is a rather debatable issue, though.<sup>30</sup>

One further note: the main arena for the debate concerning the relationship of democracy and well-being, social choice theory, has been dominated by a theory labelled 'revealed preference theory'. It is sometimes discussed under the heading of a desire- or preference-satisfaction theory of well-being.<sup>31</sup> Its main idea is that one can observe or in other ways assess the actual or hypothetical choices of an individual among various options and from

---

<sup>27</sup> Nozick (1974: 42f.). Cf. even Griffin's simulation objection (1988: chapter 1).

<sup>28</sup> See e.g. Crisp (2006: 117ff.), Sumner (1999: chapter 4).

<sup>29</sup> See e.g. Griffin (1988), Brandt (1998) and Parfit (1984).

<sup>30</sup> Tännsjö (1998: 89–91) calls this 'the argument from explanatory impotence'. Cf. Crisp (2006) and Sen (1985). Griffin (1988: ch.2) attempts to reconcile both claims, while holding on to the priority of desire.

<sup>31</sup> See e.g. Sumner (1999: chapter 5); Tännsjö (1998: chapter 6).

these choices derive an individual welfare function. Given certain information- and rationality assumptions, one can then expect the individual to behave according to her welfare function, that is, *as if* she would set out to maximise her welfare. However, this is a descriptive theory, which defines ‘welfare’ in accordance with its purpose as an analytical tool to predict behaviour. As a theory of well-being, it is relevant for this study only to the extent that it is also a plausible or correct evaluative theory about what is good for people.<sup>32</sup>

**Objective list theories.** Theories of well-being are usually called ‘objective list theories’ whenever they list a number of items that do not include, or that go beyond, the items of pleasure (and pain), or desire fulfilment. Such items may be, e.g. health, autonomy, knowledge, friendship, wealth, respect and the like. Substantial objective list accounts claim that a number of these items simply are the constituents of well-being. All they have in common is that they are — allegedly — good for people. A formal or explanatory objective list theory rather says something about what makes such things good. One suggestion is that they perfect our human nature.<sup>33</sup> Another is that they would be chosen by competent judges.<sup>34</sup>

Some object that objective list accounts do not take an individual's own priorities seriously. If one does not care for knowledge, why should (more) knowledge make life better for oneself? One answer strategy is to include the item of autonomy on the list, thus giving individuals’ priorities greater (if not ultimate) weight. Another strategy is to list capabilities that are necessary to enable individuals to ‘function’, but consider it being up to them to determine which functionings to realise in their lives.<sup>35</sup> Likewise, the list may be made up of ‘primary goods’, which are necessary for realising a wide range of life plans, according to the individuals’ own choices.<sup>36</sup> It is debatable whether these answers can make sense of the idea that what makes them well off matters to people — or indeed whether they should make sense of it.

Not every philosophical theory about well-being falls neatly into one of the above approaches.<sup>37</sup> And there are other, more complex, proposals on how to

---

<sup>32</sup> Some proponents of revealed preferentialism warn of committing the ‘causal utility fallacy, which says that decision-makers choose *a* over *b* because the utility of *a* exceeds that of *b*’ (Binmore 2007: 4). For a good discussion of instrumentalism in economics, see Rosenberg (2008: 90–95).

<sup>33</sup> See Hurka (1993).

<sup>34</sup> Some would interpret John Stuart Mill (1998) as a proponent of this view; this is, however, debatable.

<sup>35</sup> See Nussbaum(2000).

<sup>36</sup> See Rawls (1999: 54f.).

<sup>37</sup> Sumner characterises his theory, according to which well-being consists of both happiness and a fulfilment of a desire-like attitude (‘endorsement’), as ‘something in between’ hedonis-

classify these theories.<sup>38</sup> Anyhow, most of these theories seem to have a common core, concerning the things that in the end generally constitute or contribute to our well-being: health; education; close relationships; freedom of expression, occupation, assembly and religion; income or wealth; respect and so on. These things seem to be what is (intrinsically or instrumentally) good for us, even from a non-philosophical, common-sense perspective. So even though there is a lot of theoretical disagreement about the correct or most plausible philosophical framework, there seems to be a significant overlapping consensus between philosophers and laymen about what in general makes people well off. This means that studying a decision rule from a well-being perspective should be a broadly intelligible endeavour for most people.

The above theories advise us on how to compare given options, referring to possible states of affairs, with regard to how good they are for a particular individual. Roughly, the more pleasure, desire-satisfaction or objective-list items there are for an individual in a possible state of affairs, the better the option referring to this state. Arguably, the compared states of affairs should not only contain immediate consequences of adopting the option, but also more remote effects. This allows us to say, correctly I think, that keeping today's dentist appointment is better for me than skipping it, although it induces pain, or frustrates my desire to avoid dentists. The contributive value of the option 'keeping today's dentist appointment' to my well-being must plausibly be determined from the longer perspective of how well off it makes me over my entire life, or at least over a longer stretch of time. (Sometimes we explicitly ignore such a longer perspective and say things as 'Right now, skipping today's appointment would be so much better for me'. I disregard such limited claims concerning well-being in this study.)

### 2.3.2 The common good

Now that we have a better grip on possible interpretations of individual well-being, how does this notion relate to the common good?

There are (at least) two competing notions of the common good that have to be carefully distinguished. One is what we may call the *aggregative common good*. This is simply the aggregate (in some form) of the well-being of the individuals constituting the group. There are also *non-aggregative* accounts, that is, accounts that do not conceive of the common good as such an aggregate. The main thesis of my study refers to the former. I return to it

---

tic and desire theories (Sumner 1999: 175). Griffin (1988) is claimed to propose a desire-fulfilment theory according to some and an objective list theory by others. The hedonist base of Mill's (1998) version of utilitarianism has been argued to be closer to an objective list version.

<sup>38</sup> See e.g. Heathwood (2010) and Haybron (2008).

after briefly characterising the latter. How should we understand the concept of a non-aggregative common good?

There is a sense of the common good in which we sometimes say that something is good or bad *for a group*, independent of the group members' individual well-being. We say things such as: 'She did what was best for her football team' or 'Things were going well for the research group when they discovered a new species' or 'The battalion was badly off when it run out of ammunition' or the like. What is good for the group here is derived from a goal, end, or function, which the group is taken to serve. Arguably, football teams have the goal of winning (a game, the league, a championship), research groups have goal of arriving at new knowledge, battalions have their function within the chain of command and master plan for winning the war, and so on. This, then, is a teleological interpretation of the common good. As it is derived from a group's function, it is wholly independent of its members' individual well-being. Of course, it may also be in the individual football players' or scientists' self-interest to win the game or arrive at new knowledge, but such self-interest (or the possible lack thereof) is not constitutive of the groups' goal or the common interest in this sense. This means that some measure that makes a group better off may decrease every group member's individual well-being, and vice versa. For illustration, just imagine a battalion at war winning a crucial victory that leaves all of its members severely injured and traumatised, or a research team that abandons a scientifically promising, but personally hazardous excursion.

This teleological sense of the common good could possibly be extended to less functionally defined groups, such as families or nations. This could be done by claiming them to be subjected to a moral imperative within some normative framework and would hence rely on some moral interpretation of the common good. What is good for the group could then be defined in terms of such a moral end.

In this study, I am not concerned with the common good in any non-aggregative (e.g. teleological or moral) sense. Rather, I propose that we understand it in an aggregative (individualist) manner. This means that what is good for a group is derived from the well-being of its individual members; that the common interest maps on all of their individual interests.

A further question is then how this derivation or mapping is conducted: is the common good simply the sum of its members' well-being? Or do other things matter, such as the distribution of well-being, or even further values?

There are different theories about the common good. These theories provide different answers by identifying patterns of the group members' well-being and ranking these patterns according to certain criteria. We can refer to these as *axiological criteria* of the common good. Such a ranking then allows us to say that some pattern *A* is collectively better (or worse, or equally good as) some pattern *B*. Some of them are *welfarist* theories that state that only individual well-being contributes to the common good, even though

they differ on exactly how it contributes. Others are *non-welfarist* theories, stating that there are values other than individual well-being that contribute to the common good. (I disregard anti-welfarist theories that state that individual well-being is not relevant to the common good at all.) Let us look at some important theories and characterise them by stating the respective axiological criterion by which they rank patterns of individual well-being, starting with a number of welfarist criteria.

***The sum-total criterion.*** This criterion ranks the options according to the sum-total of the amount of individual well-being across all group members.<sup>39</sup> This criterion is insensitive to the distribution of well-being across a group. If one option gives you an enormous amount of well-being, and none to me, it is ranked above its alternative that gives each of us barely mediocre amounts. Other welfarist criteria, however, take the distribution into account when ranking the options. They do so by weighting the contributive value of individual well-being towards the common good according to certain considerations.

***Prioritarian criteria.*** This is a family of criteria that gives priority to the worse off, by assigning greater weights to their well-being than to that of the better off.<sup>40</sup> Depending on how these weights are assigned, different prioritarian criteria — giving more or less weight to the worse off — can be distinguished. In general, the collective value of an option is determined by summing the weighted amounts of individual well-being within the group generated by this option, with weights chosen in proportion to each individual's overall well-being (prior to or in the absence of the option's realisation). Thus, if realising one option rather than another would give you and me the same rise in well-being, yet you are overall far worse off than me, then your potential gain will have a much higher contributive value than mine.

---

<sup>39</sup> This criterion is part and parcel of e.g. total utilitarianism. In Sen's (1985: 175) words, it is the conjunction of two claims: '*welfarism* (the goodness of a state of affairs is given by the goodness of the utility information regarding that state)' and '*sum ranking* (the goodness of utility information is given by the sum-total of the utilities in question)'. The following prioritarian, sufficientarian, maximin and leximin criteria reject the *sum ranking* claim, while the non-welfarist criteria reject the *welfarism* claim. Note that I do not consider an *average criterion* since I only consider decisions with options that generate distributions of well-being across the same — and thus same-sized — group of individuals, and since for same-number groups, the average criterion is extensionally equivalent with the sum-total criterion.

<sup>40</sup> See e.g. Arneson (2000: 343): 'the [contributive] value of obtaining a benefit (avoiding a loss) for a person is greater, the greater the well-being gain that the person would get from it (the smaller the loss in well-being), and greater, the lower the person's lifetime expectation of well-being prior to receipt of the benefit (loss)'.

**Sufficientarian criteria.** This family of criteria gives exclusive priority to the well-being of individuals below a certain threshold.<sup>41</sup> This can be done in several ways. It can be done by counting the number of individuals above a given threshold of well-being and ranking the options in accordance with this number. Thus an option that lifts one more individual above the threshold than its alternative is better. This has been called a ‘*Headcount*’ sufficientarian criterion. Or one could consider exclusively the options’ distribution of well-being for all and only those individuals who are below the given threshold. Then, an option is better than its alternative if, e.g., those individuals’ sum-total of well-being (or sum-total of prioritarian-weighted well-being<sup>42</sup>) is greater. This has been called a ‘*Negative Thesis*’ sufficientarian criterion.<sup>43</sup>

**The maximin and leximin criteria.** A maximin criterion ranks distributions of individual well-being according to the levels of well-being of the worst off person(s) in each distribution.<sup>44</sup> So if I am the worst off individual — at a catastrophically low level — with one given option and you are the worst off — yet at a rather decent level — with its alternative, then the former is worse than the latter. A leximin criterion can rank an option as better than another even if both make the worst off equally badly off, by considering those who are next worst off — and if they are equally badly off, the next to next worst off, and so on.<sup>45</sup> Going from the bottom and upwards, when a difference in the levels of well-being between the compared individuals turns up, the option that makes these individuals better off is better than its alternative.

**Non-welfarist criteria.** Apart from the above welfarist criteria, there are criteria that are sensitive not only to individual well-being and its distribution, but also other alleged values.<sup>46</sup> This may, for instance, be a *desert-*

---

<sup>41</sup> See e.g. Frankfurt (1987: 31) ‘[One] response to scarcity is to distribute the available resources in such a way that as many people as possible have enough or, in other words, to maximize the incidence of sufficiency’. For a critical review of different versions of sufficientarianism, see Casal (2007).

<sup>42</sup> See e.g. Crisp (2003: 758) who states that ‘absolute priority is to be given to benefits to those below the threshold [...]. Below the threshold, benefiting people matters more the worse off those people are, the more of those people there are, and the greater the size of the benefit in question’.

<sup>43</sup> For these labels, and further versions of sufficientarianism, see Shields (2012: 103).

<sup>44</sup> Cf. Rawls (1999: 266): ‘Social and economic inequalities are to be arranged so that they are [...] to the greatest benefit of the least advantaged [...]’. My interpretation of the maximin criterion is in accordance with Broome’s proposal to understand ‘the least advantaged’ or ‘the worst off’ *de dicto* rather than *de re* (Broome’s note in Parfit 1984: 492).

<sup>45</sup> See e.g. Sen (1970: chapter 9)

<sup>46</sup> Such an incorporation of ‘non-welfarist considerations’ is mentioned, but not further developed, by Fleurbaey (*mimeo*: 6).

*sensitive criterion* according to which the contributive value of an individual's well-being is the sum of her well-being level and the degree of the 'fit' between this level and the level she would deserve.<sup>47</sup> So if one option makes me much better off than you, yet I deserve much less well-being (say, because I wronged you and ought to be punished for it), while you get exactly what you deserve, the contributive value from your situation may be greater than mine. In a similar manner, other alleged values such as *liberty* or *equality* could be integrated into criteria of the common good.

Such an integration of other alleged values may also be achieved by constraining which distributions of individual well-being can be collectively ranked at all. One such constraint may be that only those options that do not imply violations of people's *autonomy* or *basic rights* can be ranked by a criterion of the common good. Those that do would then constitute a separate class of rights- or autonomy-violating options (which then may be classed as impermissible). So, an *autonomy-constrained criterion* could rank the options according to, e.g. their sum-totals of individual well-being, under the constraint that they may not violate anyone's autonomy. Take as a simple example a case inspired by one of Nozick's examples: a bachelorette considers which one of her suitors she should marry.<sup>48</sup> Consider just three options: the bachelorette marries Adam, marries Beth or marries none at all. Say that the first option comprises most well-being in total: Adam has a large family and they all would be best off if the two got married, and this would far outweigh the bachelorette's and Beth's lack of well-being. The second option would comprise the second most well-being — even though Adam and his family would now lack the high levels of well-being, the bachelorette and Beth would be quite well off. The third option is, say, equally bad for all three of them. On a sum-total criterion one would then have to say that marrying Adam is collectively best, that is, in the common interest. Yet an autonomy-constrained criterion would arguably mark this option as forbidden. Being forced to marry against one's will clearly violates a person's autonomy, while being forced to *refrain* from marrying someone does not. Marrying Adam is then classed as an autonomy-violating and thus forbidden option which thus cannot be in the common interest. Among the remaining two options, marrying Beth is collectively better than marrying none at all. Since

---

<sup>47</sup> See e.g. Feldman (1995: 195): 'In general, the idea is that deserved goods make the world much better; undeserved goods do not make the world better. Deserved evils do not make the world worse; undeserved evils make the world much worse. If we focus on these "desert-adjusted values" when we rank possible worlds, we have a way to incorporate considerations of justice into our axiology'. My interpretation of Feldman's account in terms of a 'fit-idea' is in accordance with Arrhenius' (2006: 2) proposal that the contributive value 'is determined by the sum of the value of pleasure and the value of the fit between pleasure and the recipient's desert'.

<sup>48</sup> Cf. Nozick (1974: 269).

it does not violate anyone's autonomy, the criterion selects it as the collectively best option.

Stating these criteria is a rather simple matter. Using them to collectively rank different options, even in a heavily idealised context as the present one, carries with it a number of controversial presuppositions. All of these criteria require that we can meaningfully compare individual well-being across different persons. That is, they require that well-being, as specified by the relevant theory, be interpersonally comparable.

For instance, in order to make the claim that 'Option  $x$  is collectively better than option  $y$ ', the *sum-total* criterion presupposes that we can, for each individual, 1) rank the options according to her well-being levels, and 2) count the units of her well-being differential between the options; that we can, within the group, 3) meaningfully compare these differentials among all individuals; and for each option, 4) add the units by which it is higher-ranked than each of its alternatives by any of the group's individuals. The ranking of the options is then in accordance with the number of units we added up for each of them. In order to 'meaningfully' compare well-being differentials between individuals, the units have to be measured on an interval scale — a scale with arbitrary origins, on which 'the numerical value on one of the scales is transformed into a value on the other by means of an equation of the form  $x' = ax + b$ '.<sup>49</sup> The same applies to *non-welfarist* criteria that maximise the sum-total of well-being within the given constraints.

*Prioritarian* and *negative-thesis sufficientarian* criteria also rely on these claims, but in addition presuppose that we can identify the dimension along which well-being differentials are to be discounted (in accordance with the prioritarian weights towards the worse off) or the threshold of sufficiency above which well-being differentials should be disregarded.

The *headcount sufficientarian* criterion can do without the above four claims, since it does not presuppose comparing well-being differentials at all. Instead, this criterion presupposes a threshold such that we can 1) identify all individuals who are below the threshold level of well-being with one option and above this level with the other and 2) compare the number of individuals who are above this threshold with one option with the number who are above the level with the other option. The ranking of the options is then in accordance with these numbers.

Likewise, the *maximin* and *leximin* criteria are not concerned with well-being differentials. They instead presuppose that we can 1) identify the worst

---

<sup>49</sup> Stevens (1946: 679). Note that when the number of people is fixed, as here assumed, we can compare the sum-total of two options even without knowing things such as 'Abby is twice as well off as Beth'. The latter comparison requires measurement on a ratio scale — a scale with a non-arbitrary origin, on which the numerical value is transformed between scales by means of an equation of the form  $x' = ax$ .

off individual(s) and then 2) ordinally compare their levels of well-being with each other. The ranking of the options is then in accordance with these latter levels.

I have now proposed a variety of criteria of the common good and stated what each requires from a theory of well-being, in terms of measurability and comparability. As we will see in 3.2 and 3.3 below, the weighted majority rule can be shown to be collectively optimal for all of the stated criteria. In the meantime, I presuppose a sum-total criterion for the sake of simplicity (unless stated otherwise).

### 2.3.3 Terminology

I use the term (*individual*) *well-being* as referring to what is good for an individual. *Self-interest* is used in the following way. What promotes (maximises) one's well-being is in one's self-interest. An action that promotes one's well-being is an action in, or according to, one's self-interest. An action that one perceives promotes one's well-being is an action in, or according to, one's perceived self-interest. Someone who votes according to her self-interest is called a *self-interest voter*. A voter who is motivated by her (perceived) self-interest in her voting behaviour is called a *self-interested voter*.

The term *common good* is used to refer to the collective aggregate of individual well-being according to the relevant criterion of the common good. *Common interest* is used in the following way. What promotes (maximises) the common good is in the common interest. An option that is in the common interest is called the *common-interest option* or the *collectively best option*. An action that promotes the common good is an action in, or according to, the common interest. An action that one perceives to promote the common good is an action in, or according to, the perceived common interest. Someone who votes according to the common interest is called a *common-interest voter*. A voter who is motivated by (her perception of) the common interest in her voting behaviour is called a *common-interested voter*.

## 2.4 Democracy and the promotion of the common good

The main thesis of the present study concerns a number of arguments from collective optimality, which claim to establish that deciding by weighted majority rule promotes the common good (given certain assumptions). During the last sixty years, a number of good objections to similar claims have been advanced. They have not been made specifically against arguments concerning the *weighted* majority rule, but some of them have haunted versions of the argument in favour of closely related democratic decision methods. A number of these objections, concerning the *simple* majority rule, have

already been considered in the initial paragraphs of this introduction: the problems of the cyclical and the tyrannical majority (and to some extent the problem of the gerrymandered majority). As was claimed— and as has yet to be shown in detail — the *weighted* majority rule escapes these objections.

However, there is an influential objection to arguments from collective optimality *in general*. This objection is known as *Arrow's impossibility theorem*. Some take it to establish that *no* democratic decision rule can be collectively optimal. Is any argument from collective optimality thus doomed from the outset? It is not. To see why, let us briefly consider Arrow's theorem.

#### 2.4.1 Arrow's theorem and related problems

According to Arrow's theorem, whenever there are three or more options, and (a finite number of) two or more voters, there is no *social welfare function* that can satisfy a set of apparently reasonable conditions. A social welfare function is an aggregation of individual rankings of the options into a collective ranking that is complete and transitive. Arrow shows that there is no such function that satisfies four apparently reasonable conditions: it generates a collective ranking for all possible profiles of individual rankings (*Universal Domain*) that is not identical with some single individual's rankings (*Non-Dictatorship*), while collectively ranking an option  $x$  above an option  $y$  if every individual ranks  $x$  above  $y$  (*Pareto Efficiency*) and deriving the collective ranking of  $x$  and  $y$  exclusively from individual rankings of  $x$  and  $y$  (*Independence*).<sup>50</sup>

Simple majority rule does satisfy Arrow's four conditions.<sup>51</sup> Still, it lies outside the scope of Arrow's theorem since its straightforward application is limited to binary decisions. (This of course limits its general practical relevance as well.) Simple majority rule can be extended to work for decisions with more than two options. The *Condorcet rule*, e.g., first applies simple majority rule to all possible pairs of options. Second, from the resulting binary rankings, it constructs a collective ranking of all the options according to an ordinal ranking rule: an option  $x$  is collectively ranked above an option  $y$  if and only if  $x$  receives more votes than  $y$  in a pairwise vote. (This is the decision procedure hinted at in the initial — hypothetical — French presidential election case. Of course, the French do not actually use the Condor-

---

<sup>50</sup> Arrow (1963). For a simple and elegant proof of Arrow's theorem, see Sen (1995: 4).

<sup>51</sup> Simple majority moreover uniquely satisfies three further appealing conditions: *Decisiveness* (for any profile of individual rankings, the decision rule specifies a unique collective ranking — allowing ties), *Monotonicity* (if, ceteris paribus, one or more voters change their individual rankings such that option  $x$  is ranked higher than before, then, if the collective ranking is changed, it ranks  $x$  higher than before), *Anonymity* (if, ceteris paribus, any individual rankings are assigned to different voters, the collective ranking does not change), and *Neutrality* (if, ceteris paribus, the labels of the options are changed, the collective ranking of the options does not change); see May (1952).

cet rule when voting for president.) The problem with the Condorcet rule, as we have already seen, is that it may construct cyclical collective rankings (e.g. ranking Hollande over Sarkozy over Le Pen over Hollande).<sup>52</sup>

It can be objected that democratic decision methods, such as the simple majority rule or its Condorcet extension, should not be considered social welfare functions. They aim to derive a collective choice from individual input rather than a collective ranking. And even if they would base the choice on such a ranking, it is not obvious why this ranking would have to be complete and transitive.<sup>53</sup> The question is then whether there can be a *social choice function* that simply selects a non-empty set of ‘winning’ options and that satisfies Arrow's four conditions. Alas, this is not the case. Amartya Sen's ‘general choice functional impossibility theorem’ shows that, once the four conditions are adapted to this social choice context, their conjunction may lead to empty choice sets.<sup>54</sup>

This clearly spells trouble for democratic decision methods, understood as social choice functions, if they are to satisfy Arrow's four conditions. According to William Riker, Arrow's (and related) theoretical results can be shown to lead to immense and ubiquitous problems in everyday practical democratic decision-making.<sup>55</sup>

The legacy of Arrow and Riker has inclined some political theorists to consider the link between democracy and welfare to be severed beyond repair — in theory as well as in praxis. Others however maintain that, despite these results, democratic decision-making can be argued to be welfare-enhancing. Gerry Mackie argues forcefully and in profound detail that Riker's (as well as others') ‘irrationalist interpretations of social choice theory are based on unrealistic assumptions, or illustrate logical possibilities rather than empirical probabilities, or emphasize remediable problems, or are outright mistaken’.<sup>56</sup> James Buchanan and Gordon Tullock recognise the economic dimension of voting and assume that the individual voter may rationally engage in vote-trading (logrolling), and they show that this form of ‘manipulation’ has beneficial effects on the group's welfare under certain conditions.<sup>57</sup> Jonathan Riley argues that simple majority rule approximates collective optimality ‘when interpersonal utility comparisons are impractical or impossible’.<sup>58</sup> A theorem by Douglas Rae and Michael Taylor shows that

---

<sup>52</sup> This ‘paradox of cyclical majorities’ was first discovered by Condorcet, and later rediscovered by Charles Dodgson (Lewis Carroll) and, separately, Black (1958). The Arrow theorem is in fact a generalisation of the paradox. The actual Arrowian label ‘General Possibility Theorem’ refers to the *possibility* of such an intransitive collective ranking.

<sup>53</sup> For this criticism, see e.g. Buchanan (1954).

<sup>54</sup> Sen (1995: 6), cf. Sen (1993).

<sup>55</sup> Riker (1982).

<sup>56</sup> Mackie (2003: 16).

<sup>57</sup> Buchanan and Tullock (2004). I return to this claim in 7.1 below.

<sup>58</sup> Riley (1990: 336).

simple majority rule does maximise individual well-being under certain conditions.<sup>59</sup>

There is another strategy of refuting the pessimistic view that one cannot show democracy to promote collective welfare. The strategy is to argue that some of Arrow's conditions should be relaxed. It sees the impossibility results not as dooming the idea entirely, but rather as setting out limits within which it can be defended. As Sen rightly remarks:

‘The real issue is not, therefore, the ubiquity of impossibility (it will always lie close to the axiomatic derivation of any specific social choice rule), but the reach and reasonableness of the axioms to be used. We have to get on with the basic task of obtaining workable rules that satisfy reasonable requirements’.<sup>60</sup>

My study is conducted in the spirit of this strategy. I clarify why this is so in the next section.

## 2.4.2 Returning to the weighted majority rule

The weighted majority rule violates Arrow's Independence condition. This can be seen from the above *Dinner plans 1* and *Dinner plans 2* cases. We have assumed that each of the three voters keeps the same ranking in the first and in the second case: Abby and Beth rank *out* over *in*, while Charlie ranks *in* over *out*. Still the collective rankings differ: in *Dinner plans 1*, *out* is collectively ranked over *in*, whereas in *Dinner plans 2*, the collective ranking is reversed. This shows that the weighted majority rule derives the collective ranking *not* exclusively from information about the individual rankings, but also from information about interpersonally comparable stakes. This is a clear breach of the Independence condition.

But this is as it should be, or so one might argue. The gist of the condition is that no irrelevant information should affect the result (such as ‘third’ options that are not even up for decision). Yet the issue of how much well-being is at stake for the voters, in comparison, between the two options at hand can surely be argued to be of relevance for a democratic decision rule. It certainly is within the present context of collective optimality. Hence, it seems only reasonable that the independence condition should be relaxed.<sup>61</sup>

The above *Dinner plans* cases consider decisions with two options. What about decisions with more options? Consider the following case.

---

<sup>59</sup> Rae (1969) and Taylor (1969). I return to this theorem in the Appendix below.

<sup>60</sup> Sen (1999b: 354).

<sup>61</sup> The reason information about interpersonally comparable well-being is considered problematic is one of alleged infeasibility rather than inappropriateness: it has been thought impossible to securely attain such information. This is, however, a debatable stance. I briefly comment on this matter in 3.4 below.

**Dinner plans 3.** A group of three, Abby, Beth, and Charlie, needs to decide whether to go *out* to a restaurant, stay *in* and cook, or just *order* dinner on the phone. Let us assume that Abby's ranking is  $in > out > order$ , Beth's ranking is  $out > order > in$  and Charlie's ranking is  $order > in > out$ .<sup>62</sup> Moreover, all three are similarly affected by the options, in the sense that each gets the same amount of well-being from the same position in their ranking, say, two units from the top-ranked option, one unit from the second-best option, and zero units from the bottom option; see *Table 1*. This means that Abby's stake (well-being differential) is one unit in the decision *in* vs. *out* (her top option vs. second-best option), one unit in the decision *out* vs. *order* (her second-best option vs. bottom option) and two units in the decision *in* vs. *order* (her top vs. bottom option), and similarly for Beth and Charlie. (Note that, in fact, the exact figures in each cell are not important. What matters is that the higher they are, the better the option for the individual. Moreover, the differences between the figures give us the — interpersonally comparable — well-being differentials, and thus the stakes, for each individual.)

	<i>Abby</i>	<i>Beth</i>	<i>Charlie</i>
<i>in</i>	2	0	1
<i>out</i>	1	2	0
<i>order</i>	0	1	2

*Table 1*

The weighted majority rule, as specified above, operates on pairs of options. If applied to all possible pairs, it will assign different amounts of votes to each individual, depending on whether the decision concerns her top and bottom options (more stakes and thus more votes) or not. So in the decision *in* vs. *out*, the rule assigns one vote each to Abby and Charlie, while assigning two votes to Beth. Since Abby and Charlie vote for *in*, while Beth votes for *out*, there is a tie (denoted by '~') of two votes for *in* and two votes for *out*, such that  $in \sim out$  in the collective ranking. Similar results apply to decisions *out* vs. *order* (where one vote is assigned to Abby and Beth each and two votes to Charlie) and *in* vs. *order* (where one vote is assigned to Beth and Charlie each and two votes to Abby). The result is that all pairs tie. It could then be reasonably claimed that the collective ranking is a global tie, that is, a tie across all options:  $in \sim out \sim order$ . Then, it could be argued that this is as it should be, from a common-good perspective. As the figures in *Table 1* indicate, each option is as collectively good as the other (according to the sum-total criterion). The problem is not that there is a tie, but only that

---

<sup>62</sup> This is the typical case where simple majority rule would effect a collective cycle. Since it assigns one vote to each voter, there emerges a cycling majority that collectively ranks  $in > out$ ,  $out > order$  and  $order > in$ .

we now need a complementary tie-breaker. One intuitively plausible candidate would be an even-chance lottery among the three options.<sup>63</sup>

So, in all of the above *Dinner plans* cases, the weighted majority rule (complemented with an appropriate tie-breaker) manages to select the collectively better option (according to the sum-total criterion). It moreover arguably leads to better results in the ‘majority conundrum’ cases, stated in 1.1 above, by avoiding the problems of *gerrymandering*, *tyrannical* and *cyclical majorities*. All this merely indicates that the weighted majority rule is a quite promising collective decision rule. Further investigation is needed concerning its performance. The present study is a contribution to this project.

## 2.5 Notes on methodology and philosophical framework

I conclude this chapter with a couple of notes on the moral-philosophical framework of this study and the methodology and relevance of formal models.

### 2.5.1 The moral-philosophical framework

This study deals with a democratic decision rule — the weighted majority rule — as an ideal institutional structure, that is, a common and rather stable system of rules and constraints for (idealised) individual behaviour, providing rules and incentives to (idealised) individual agents. It thus falls within the domain of political philosophy.

The study approaches its object from a moral-philosophical perspective, according to the following line of thought. As stated above, my point of departure is a set of alternative criteria of the common good, in the light of which the performance of the weighted majority rule is to be evaluated. These criteria, as well as the related theories of individual well-being, belong to a subfield of moral philosophy, namely normative ethics. This study can therefore be described as applying normative-ethical criteria to a specific decision rule and its outcomes. Thus construed, it is an exercise in applied ethics.

Situating my study within this framework may raise questions as to how the former relates to other theories within moral philosophy. To avoid limiting the relevance of my study unnecessarily, I want to neither presuppose any specific normative theory that makes claims about what is right, what we ought to do, etc. Nor do I want to rely on any specific meta-ethical theories about the nature, accessibility, and meaning of norms or values. I merely state that, according to the outlined theories of well-being and the common good, some states of the world, or properties, or facts, are good for people

---

<sup>63</sup> I return to the tie-breaker question in 3.2 and 7.3.2 below.

and collectively good, respectively. It does not matter whether these states, properties or facts are taken to be of intrinsic or final value, or neither. My claim is that, given the relevant theories, the use of the weighted majority rule for collective decision-making is of instrumental value insofar as it promotes what they declare to be collectively good.

### 2.5.2 Formal models

This study applies a number of criteria of the common good to the output (outcomes) of the weighted majority rule, under certain assumptions, concerning voters' behaviour, which determines the input (votes). These three components — input assumptions, operation of the weighted majority rule and resulting output — constitute a formal model that allows us to isolate and examine certain details of the decision rule. As Jane Mansbridge succinctly puts it:

‘Models help us think by making it possible to reduce complexity with a few assumptions, then draw out the implications of those assumptions, creating a new kind of complexity we could not capture if we tried to account at the outset for all aspects of reality’.<sup>64</sup>

Similarly, in Sen's words, formal modes of reasoning provide us with ‘the opportunity to consider various results which [...] are not easily anticipated without formal reasoning. Informal insights, important as they are, cannot replace the formal investigations that are needed to examine the congruity and cogency of combinations of values and of apparently plausible demands’.<sup>65</sup>

Formal models are frequently used in social choice theory, which lies at the border between philosophy and economics. To the extent that these models are used to analyse and evaluate institutional structures, they are of relevance to political philosophy. To see how they are employed, let us briefly look at some prominent examples.

Arrow's already stated impossibility theorem can be understood as derived from such a model, e.g. in the following way. The input assumption is Unrestricted Domain, allowing all individual rankings of the given set of (more than two) options. Arrow then shows that there is *no* rule that can aggregate individual preference rankings (input) into collective preference rankings (output) that satisfy the criteria of Completeness and Transitivity, of Pareto Efficiency, Independence and Non-Dictatorship. The result of this model — that the output must violate this set of plausible normative criteria (Arrow calls them ‘natural conditions’)<sup>66</sup> — can be used to construct a pow-

---

<sup>64</sup> Mansbridge (1990: 254).

<sup>65</sup> Sen (1999b: 353).

<sup>66</sup> Arrow (1963: 2).

erful argument against any such rule, given the Unrestricted Domain input assumption.

With a similar model it can be shown that if the input assumption is modified, there is one rule that generates an output satisfying the same normative criteria. As Duncan Black suggests, the Unrestricted Domain assumption can be replaced by a *Single-Peaked Preference* assumption.<sup>67</sup> To sketch this latter assumption, it states that there is some single linear dimension along which all the given options can be ordered (such as the traditional left-right dimension in politics). Moreover, for every voter, if an option  $y$  is located, on that dimension, between her top-ranked option  $x$  and another option  $z$ , then she ranks  $y$  above  $z$ . (For instance, if the left-wing candidate is top-ranked by some voter, and the centre candidate is located between the left- and right-wing candidates on the left-right dimension, this voter's ranking concerning these three candidates is single-peaked if she then ranks the centre candidate above the right-wing candidate.) Arrow shows that with this input assumption, operation of the *Condorcet rule* (the pairwise application of simple majority rule) does generate an output that satisfies the above criteria.<sup>68</sup> This means that the model, with its *Single-Peaked Preference* input assumption, can be successfully employed in an argument for the Condorcet rule, in the light of these normative criteria.

Buchanan and Tullock use a different, as they say, 'extremely weak ethical criterion for "betterness"' to evaluate decision rules: unanimous consensus from rational, self-interested individuals.<sup>69</sup> The input assumption of their model is that voters *behave rationally and self-interested*. Under this assumption, they analyse the operation of unanimity rule and simple majority rule. The output of the model is defined in terms of the costs and benefits of these decision rules' outcomes (for which Buchanan and Tullock explicitly take vote trading into account), as well as the costs and benefits of employing the decision rules themselves. This output is then evaluated in the light of the stated normative criterion — whether rational, self-interested individuals would consent to it.

The above-sketches three models differ from each other, and from my subsequent model, in the specification of their components (input assumptions, operating decision rule and resulting output) as well as in their applied normative criteria. However, they share a crucial feature that tends to make people suspicious: their input assumptions are (generally) false. For instance, they assume people to have complete and transitive (or even single-peaked) preference orderings, or to be rational and self-interested. This obviously does not hold for many people in many real-life situations.

---

<sup>67</sup> Black (1958: 14–25).

<sup>68</sup> Arrow (1963).

<sup>69</sup> Buchanan and Tullock (2004: 13).

The same can be said of the formal model underlying the coming argument from collective optimality. It works on the assumption that voters always vote based on their self-interest. Yet we know from experience that voters cast their votes according to all kinds of motivations. This raises the question of whether the results we can derive with logical precision from formal modelling can ever be practically relevant.

The kind of answer we can provide depends on the purpose of the model. Consider model building in economics. The purpose of modelling in economics is usually to provide explanations and predictions for real-world states of affairs. To illustrate with a simple example: say that we observe a decrease in the demand for olive oil. We wonder: how can this decrease be explained and could it have been predicted? A possible answer starts from a traditional economic model, with the (unrealistic) input assumption that all buyers are rational, self-interested individuals, and an analysis of the operation of a general rule, such as the law of demand (stating that as the price of some product goes up, demand for the product among rational self-interested buyers will decline). If we can observe that the price for olive oil in fact has risen, the model can be employed to explain the decrease in demand and to predict such decrease (*ceteris paribus*). Here, the general idea is that models are good or useful to the extent that they make correct predictions. As far as we are interested in such predictions, these models are practically relevant — even if their assumptions may be false.<sup>70</sup>

This approach to *explanatory-predictive modelling* can also be found within political theory. Anthony Downs analyses the operation of political party-systems under the (unrealistic) input assumption that voters and representatives alike are rational, self-interested individuals.<sup>71</sup> He then devotes an appendix to comparing statistical real-world data with the predicted output of his model and finds that they match considerably. If this is right, the data can be taken to corroborate the model — they show the model to be good or useful, and thus practically relevant, even though the model's assumptions may be false.

However, in the present context of what we may call *normative modelling*, models have a different purpose. They are set up not to give us testable predictions (and explanations) of certain real-world phenomena, but rather to allow insights into the logical structure of the operation of decision rules. The general idea is that formal models help us see that the following conditional is true, that is, that the antecedent cannot be true without the consequent being true as well:

If [input assumption] and [operation of the decision rule], then [generated output].

---

<sup>70</sup> Cf. e.g. Friedman (1994). For a critical discussion of Friedman's position, see Sen (1997).

<sup>71</sup> Downs (1957).

Such an insight can then be invoked in a normative argument for or against a specific decision rule. For instance, if the generated output is valuable in the light of a certain normative criterion, then we can argue from the above conditional that the operation of the decision rule is instrumentally valuable, given that the input assumption holds. Such an argument is practically relevant for all real-life cases in which the input assumption is true. That is, for all actual cases of collective decision-making where the input assumption is satisfied, we have an argument in favour of actual (correct) implementation of the decision rule in question. (Note that this may be a *prima facie* conclusion, as it may turn out that there are other considerations that speak against such implementation.)

However, immediate practical relevance is not the only or ultimate measure of what makes an argument worth our while. Reasoning within idealised conditions can bring theoretical insights that are of interest even if their direct practical relevance is has not (yet) been established. For instance, Arrow's arguments have improved our understanding of collective decision-making, even though they are based on highly idealised assumptions.<sup>72</sup> Their practical relevance for real-life decision-making has then been a further (and, to say the least, heavily debated) question.<sup>73</sup>

The main part of this study is concerned with the proposed formal models and resulting arguments from collective optimality, within idealised settings. For instance, it is shown that, *given an idealised input assumption* of self-interested voting, the operation of the weighted majority rule — specified in a variety of ways — generates different kinds of outputs, which satisfy a variety of criteria of the common good. This gives rise to a collective optimality argument for the weighted majority rule, given the input assumption (Chapter 3). It is then shown that the input assumption can be weakened, allowing even some forms of non-self-interested voting (Chapters 4 through 6). To be sure, even the emerging weakened assumptions are highly idealised. The question of whether they hold in actual (real-life) decision-making cases — that is, whether (and when) the argument from collective optimality is practically relevant — receives some speculative treatment in my final discussion in Chapter 8, but is mostly left for another occasion.

One last note on my use of formal models: I hope it does not deter the reader. Basically nothing more than elementary algebra is required. I attempt to keep things as simple as possible and summarise all the results of formal reasoning in informal language.

---

<sup>72</sup> Arrow (1963).

<sup>73</sup> Cf. e.g. Riker (1982), Mackie (2003).



## 3 A case for the weighted majority rule

### 3.1 Introduction

Let us take a look at the case for the weighted majority rule then, as proposed by Brighthouse and Fleurbaey and by Fleurbaey, respectively.<sup>1</sup> In section 3.2, I reconstruct and start to analyse their argument, which I call the *original argument from collective optimality*.<sup>2</sup> The argument shows that, given certain assumptions, the weighted majority rule is collectively optimal, according to two different criteria of the common good. I then show that by varying the assumptions slightly, a larger variety of such criteria can be accommodated. In 3.3, I state a more general *generic argument from collective optimality*, which works for any of these criteria. Section 3.4 brings out some important limitations and clarifications concerning the considered arguments. Section 3.5 concludes.

Note that, for now, the weighted majority rule is assumed to apply only to binary decisions, that is, decisions with two options. In Chapter 6, I relax this limitation in scope and assess the implications of applying the weighted majority rule to decisions with three and more options.

### 3.2 The original argument from collective optimality

Fleurbaey presents a version of the argument from collective optimality in his recent article ‘One Stake One Vote’. I start by quoting the main parts:

---

<sup>1</sup> Brighthouse and Fleurbaey (2010), Fleurbaey (*mimeo*).

<sup>2</sup> Brighthouse and Fleurbaey (2010: 138) write that the weighted majority rule instantiates what they call the ‘principle of proportionality’: ‘*Power in any decision-making process should be proportional to individual stakes*’. This principle replaces the orthodox democratic slogan of ‘one person, one vote’, with, so to speak, the proportionality slogan of ‘one stake, one vote’ (cf. Fleurbaey *mimeo*). In Brighthouse and Fleurbaey (2010) the normative justification for the proportionality principle rests on the values of equality, autonomy and on a consequentialist, *prioritarian*, conception of the common good. In my study, I ignore the principle of proportionality and instead focus exclusively on the weighted majority rule. Moreover, I focus mainly on a consequentialist justification in terms of the common good.

‘Suppose we are interested in maximizing social welfare, understood as the sum of individual utilities. [... Extending] the observation that the simple majority rule maximizes the sum of utilities in the case of equal intensity [it can be shown that] *the weighted majority rule maximizes the sum of utilities when voters have unequal intensities and their weights are proportional to the utility differences  $\Delta_i$* . It is assumed that each voter votes for the option which is more favorable to him [...].

**Proposition 1** *If the weights  $w_i$  are proportional to  $\Delta_i$ , the weighted majority rule selects the decision which yields the greater sum of utilities.*<sup>3</sup>

Fleurbaey then provides a formal proof for Proposition 1, using some simple technical notation. I now want to reconstruct Fleurbaey's main idea in a hopefully clearer and non-technical fashion. My aim is to show, first, how Proposition 1 and the ensuing proof can be interpreted in terms of a formal model, generating insight into the working of the weighted majority rule. Second, I want to show how this model can then be employed in an argument for the collective optimality of this rule, in the light of a welfarist sum-total criterion.

First, recall our definition of the weighted majority rule (see 2.2 above).

**The Weighted Majority Rule:** For all individuals and any decision with two options, (a) every individual is assigned a number of votes in proportion to her stakes, and (b) the option that receives a majority of votes is selected as outcome.

Second, let us define an individual's stakes as the differential in well-being between the two options, for this individual. To illustrate: if one of the options makes you very well off and the other miserable, then according to this claim you have greater stakes than, say, me, who is made almost equally well off by both options. More precisely, if the difference in well-being units for you is 10 times the difference for me, this means that your stakes are 10 times greater than mine. (According to the Weighted Majority Rule, you will then receive 10 times the number of votes I receive.) This also implies that someone who is equally well (or badly) off with either option has no stakes (and thus will receive no votes).

**Welfare Stakes:** An individual's stakes in a binary decision consist in her differential in well-being between the two options.

Third, we specify the voting behaviour that generates the input for the operation of the rule.

**Self-Interested Voting:** Every voter (that is, individual who has been assigned a positive number of votes) votes according to her self-interest.

---

<sup>3</sup> Fleurbaey (*mimeo*: 6).

Now we can restate the original theorem as follows.

**The Original Theorem:** For all individuals and any decision with two options, given the Weighted Majority Rule, Welfare Stakes and Self-Interested Voting, the option with the greater sum-total of well-being is selected as the outcome.

It can now be shown that the Original Theorem must be true. That is, for the specified set of cases, it can be shown that if the theorem's three conditions are satisfied, its result (regarding the selected outcome) necessarily holds as well, according to the following line of reasoning.<sup>4</sup>

Assume that the two options in the given decision are  $x$  and  $y$ , and assume that  $x$  is the option with a greater sum-total of well-being. This tells us something about the distribution of well-being differentials between the options, across all individuals. It tells us that the individuals who are better off with  $x$  must together have a greater sum of well-being differentials than those who are better off with  $y$ . (Were this not so,  $x$  could *not* yield a greater sum-total of well-being than  $y$ .) Since according to Welfare Stakes, stakes are defined as well-being differentials this simply means that the sum of stakes held by those who are better off with  $x$  is greater than the sum of stakes held by those who are better off with  $y$ . Since according to the Weighted Majority Rule, votes are assigned in proportion to stakes, the sum of votes assigned to those who are better off with  $x$  must be greater than the sum of votes assigned to those who are better off with  $y$ . Since, according to Self-Interested Voting, those who are better off with  $x$  cast their votes for  $x$ , and those who are better off with  $y$  cast their votes for  $y$ , this implies that  $x$  receives more votes than  $y$  — that is, a majority of votes (since the theorem is limited to decision with two options). Then, according to the Weighted Majority Rule,  $x$  is selected as the outcome. By assumption,  $x$  is the option with a greater sum-total of well-being. Thus, we have derived the result of the theorem from its conditions and the assumption that  $x$  is the option with a greater sum-total of well-being. The same line of reasoning can be employed if we instead assume that  $y$  is the option with a greater sum-total of well-being.

What if  $x$  yields the same sum-total of welfare as  $y$ ? Following the stated line of reasoning, in such a case the sum of votes assigned to those who are better off with  $x$  *equals* the sum of votes assigned to those who are better off with  $y$ . Then, if everyone votes for whatever option makes her better off, the weighted majority rule will not select any of the options (since none receives a majority of votes). This means that there is a tie. However, in this case it does not matter, from a common good perspective, which option is selected.

---

<sup>4</sup> The theorem follows the simple pattern of the formal model as sketched in the previous chapter (2.5.2 above): If [input assumption] and [operation of the decision rule], then [generated output]. I call what is 'given' (corresponding to the conditional's antecedent) the theorem's *conditions*, and the claim concerning the 'generated output' (corresponding to the conditional's consequent) the theorem's *result*.

So any tie-breaker could be used for this purpose. This may, for instance, be a *random* tie-breaker, which breaks the tie by means of a fair coin toss, or a *democratic* tie-breaker, which lets a democratically elected chairperson select the winning option, or the like.<sup>5</sup>

Note also that this argument ignores those who have a zero well-being differential in the decision between  $x$  and  $y$ . Since they hold no stakes and thus receive no votes, they are not relevant for the operation of the weighted majority rule. Further, since their levels of well-being for  $x$  and  $y$  make no difference for the relative sum-total for both options, they are not relevant for the value of the outcome of the decision. They can thus be safely ignored by the argument.

The Original Theorem is thus shown to be true. It can now be employed as a premise within an argument for the collective optimality of the weighted majority rule, according to a sum-total criterion of the common good. We can call this the *original argument from collective optimality*.

(1) *The Original Theorem*: For all individuals and any decision with two options, given the Weighted Majority Rule, Welfare Stakes, and Self-Interested Voting, the option with the greater sum-total of well-being is selected as the outcome.

(2) *The Sum-Total Criterion*: The option with the greatest sum-total of well-being is collectively optimal.

(3) A rule that selects the collectively optimal option, among all the given alternatives, is collectively optimal.

Hence:

(4) For all individuals and any decision with two options, given Welfare Stakes and Self-Interested Voting, the Weighted Majority Rule is collectively optimal, according to the sum-total criterion.

Note that the argument only considers the rule's performance regarding the generated outcome. Yet there may be procedural 'side-effects' from employing the rule, which affect its overall collective optimality. These side-effects are here disregarded. (I briefly return to this issue in 3.4 below.)

We have now a precise argument for the collective optimality of the weighted majority rule, given a specific criterion of the common good: the

---

<sup>5</sup> Of course, there may be additional considerations in favour of certain tie-breakers such as the above, and against others, such as that the strongest or richest or prettiest gets to decide whenever options tie. Cf. Arrhenius (2011: 20) who proposes that 'we should require of a democratic method that it doesn't put any voter in a favoured position — a so-called anonymous method — to pick the winner. Examples of such methods are a democratically elected chairman or a random device.' Such considerations may specifically apply from a common-good perspective.

sum-total criterion. Still, recall (from 2.3.2 above) that there are alternative criteria (which some might find more plausible). Can the argument be adapted to them? Fleurbaey (*mimeo*) proposes that this could be done by reinterpreting the notion of individual well-being — or in his terms, of individual utility:

‘[T]he sum of individual utilities [that is to be maximized] need not represent the classical utilitarian criterion because the [individual utilities] can be interpreted in many ways. The most convenient for our purposes is to understand them as measuring the social value of *i*’s situation for the observer. This makes it possible for the welfare criterion to incorporate a priority for the worst off as well as non-welfarist considerations in the measurement of individual wellbeing. Any continuous separable social criterion, welfarist or non-welfarist, is encompassed in our approach. The term “utility” is retained only for its simplicity and in order to make the results more easily comparable with those found in the literature’.<sup>6</sup>

However, this interpretation — and the resulting obfuscation of the distinction between the individual utility and the social (or contributive) value of the individual’s situation — is not unproblematic. First of all, we still want to be able to say that, for instance, option *x* increases *i*’s individual utility (makes her better off), compared to *y*, yet does not have greater social value (say, because *i* is way above the sufficientarian threshold with either option, so her situation has no social value at all). Second, we may want to guard ourselves against equivocating on ‘utility’. The risk is illustrated in Fleurbaey’s paper, when he suggests how voting weights could be determined when there is uncertainty about stakes (across several binary decisions). Fleurbaey proposes that ‘in some cases the specific weights can be made to rely on voters’ private information’, by assigning to each voter a certain number of votes and allowing her to distribute these votes across the options of several decisions, according to the relative importance of the decisions.<sup>7</sup> This suggestion seems more plausible when a voter’s ‘private information’ is taken to concern what makes her better off, rather than what is the social value of her situation. If this is so, then, contrary to Fleurbaey’s initial proposal, we are here not dealing with ‘utility’ in the social-value sense, but are sliding back into the ‘what makes someone better off’-sense.

There is another proposal for showing the weighted majority rule’s collective optimality. This proposal can be used as a guide on how to adapt the argument to a variety of criteria, as shown in the following section.

---

<sup>6</sup> Fleurbaey (*mimeo*: 6).

<sup>7</sup> Fleurbaey (*mimeo*: 19).

### 3.2.1 Alternative versions of the argument

Brighouse and Fleurbaey (2010) consider the weighted majority rule in the light of a *prioritarian* criterion and propose the following theorem and supporting argument:

‘Consider a prioritarian criterion maximizing the sum of  $f(I_i(x))$  over all  $i$ , where  $f$  is a concave function and  $I_i(x)$  measures the situation of individual  $i$ 's interests with option  $x$ . Suppose that the options are ranked by application of the weighted majority rule over every pair of options, individual weights being specific to every pair of options and being computed as the absolute value of the difference in  $f$ -transformed individual interests between the two options in a pair (i.e.  $|f(I_i(x)) - f(I_i(y))|$  is the weight of individual  $i$  for application of the majority rule to the decision between  $x$  and  $y$ ). Then, assuming that every individual always votes according to his interests, the options are ranked in agreement with the prioritarian criterion.

[...] The theorem itself can be explained as follows. The option which wins is such that the sum-total of weights of those who vote for it is greater than the sum-total of weights of those who vote against. Since weights equal stakes, and stakes equal the differences in  $f$ -transformed individual interests, this directly implies that the sum of  $f$ -transformed interest differences is greater for those who gain with this option (compared to the alternative option) than for those who lose. And this is equivalent to saying that the sum of  $f$ -transformed individual interests is greater with the winning option’.<sup>8</sup>

What is meant by the expression ‘ $f$ -transformed individual interests’? In the present terminology, it can be characterised as follows. Take, first, the level of well-being of each option for each person. Second, transform these levels with a function  $f$ . The specific function is given by a prioritarian criterion, which gives priority to the worse off. Third, calculate each person's stakes as the differential of her  $f$ -transformed well-being levels. Since this function is concave, it generates greater stakes for well-being differentials of the worse off than for the same well-being differentials of the better off. This gives us an alternative interpretation of an individual's stakes.

**Priority Stakes:** An individual's stakes in a binary decision consist in her differential in  $f$ -transformed well-being between the two options.

Let us look at the following case for illustration. Assume that for the two options eating *in* and eating *out*, Abby and Beth have the same well-being differential. *in* makes both equally better off than *out*, since, say, they are both equally hungry and equally broke (and thus cannot afford to pay for a full meal at a restaurant). However, Abby is happy, healthy and has many good friends and a satisfying job. Beth, on the other hand, lacks health, friends and work and is therefore rather depressed. So while Abby's and Beth's well-being differentials are equal, Abby's well-being levels for either

---

<sup>8</sup> Brighouse & Fleurbaey (2010: 143–144).

option are higher than Beth's. When these levels are transformed by the prioritarian function  $f$ , the differential is greater for the worse off, that is, for Beth, than for better off Abby. In other words, then, Beth has higher stakes — even though she would gain as much well-being from eating *in* rather than *out* as Abby.

Brighouse and Fleurbaey's argument shows that the weighted majority rule is collectively optimal in the sense that it satisfies a *prioritarian* criterion, which ranks distributions of individual well-being giving priority (according to some specified concave function  $f$ ) to the gains and losses of the worse off.<sup>9</sup> Their main idea can be restated as the following alternative theorem.

***The Prioritarian Original Theorem:*** For all individuals and any decision with two options, given the Weighted Majority Rule, Priority Stakes, and Self-Interested Voting, the option with the greater sum-total of  $f$ -transformed well-being levels is selected as the outcome.

It can now be seen that the Prioritarian Original Theorem must be true as well. That is, for the specified set of cases, it can be shown that if the theorem's three conditions are satisfied, its result (regarding the selected outcome) necessarily holds as well, according to the same line of reasoning as for the Original Theorem above — with 'well-being' replaced by ' $f$ -transformed well-being levels'.

Assume that the two options in the given decision are  $x$  and  $y$ , and assume that  $x$  is the option with a greater sum-total of  $f$ -transformed well-being levels. This tells us something about the distribution of the differentials in  $f$ -transformed well-being levels between the options, across all individuals. It tells us that the individuals who are better off with  $x$  must together have a greater sum of differentials in  $f$ -transformed well-being levels than those who are better off with  $y$ . (Were this not so,  $x$  could *not* yield a greater sum-total of  $f$ -transformed well-being levels than  $y$ .) Since, according to Priority Stakes, stakes are defined as differentials in  $f$ -transformed well-being levels, this simply means that the sum of stakes held by those who are better off with  $x$  is greater than the sum of stakes held by those who are better off with  $y$ . And since, according to the Weighted Majority Rule, votes are assigned in proportion to stakes, the sum of votes assigned to those who are better off with  $x$  must be greater than the sum of votes assigned to those who are better off with  $y$ . Since, according to Self-Interested Voting, those who are better off with  $x$  cast their votes for  $x$ , and those who are better off with  $y$  cast their votes for  $y$ , this implies that  $x$  receives more votes than  $y$  — that is, a majority of votes (since the theorem is limited to decision with two options). Then, according to the Weighted Majority Rule,  $x$  is selected as the outcome. By

---

<sup>9</sup> Brighouse and Fleurbaey (2010).

assumption,  $x$  is the option with a greater sum-total of  $f$ -transformed well-being levels. Thus, we have derived the result of the theorem from its conditions and the assumption that  $x$  is the option with a greater sum-total of  $f$ -transformed well-being levels.

Again, the same line of reasoning can be employed if we instead assume that  $y$  is the option with a greater sum-total of  $f$ -transformed well-being levels. Again, we can disregard cases where  $x$  yields the same sum-total of  $f$ -transformed well-being levels as  $y$ . Finally, we can again disregard individuals who have a zero differential in  $f$ -transformed well-being levels (and thus zero votes) in the decision between  $x$  and  $y$ .

The Prioritarian Original Theorem is thus shown to be true. It can now be employed as a premise within what we may call the *prioritarian original argument from collective optimality*.

(1) *The Prioritarian Original Theorem*: For all individuals and any decision with two options, given the Weighted Majority Rule, Priority Stakes, and Self-Interested Voting, the option with the greater sum-total of  $f$ -transformed well-being levels is selected as the outcome.

(2) *The Prioritarian Criterion*: The option with the greatest sum-total of  $f$ -transformed well-being levels is collectively optimal.

(3) A rule that selects the collectively optimal option, among all the given alternatives, is collectively optimal.

Hence:

(4) For all individuals and any decision with two options, given Priority Stakes and Self-Interested Voting, the Weighted Majority Rule is collectively optimal, according to the prioritarian criterion.

We have now a precise argument for the collective optimality of the weighted majority rule from a prioritarian criterion as well. To obtain this argument, we have merely made substantial changes in two premises of the *original argument from collective optimality*. The Sum-Total Criterion was replaced by the Prioritarian Criterion. Moreover, Welfare Stakes was replaced by Priority Stakes.

This strategy opens up for further ways to adapt the *original argument from collective optimality* to alternative criteria of the common good. Consider for instance a *sufficientarian* criterion. If it is taken to propose maximising the number of people above a given threshold of well-being, according to the headcount claim, we can now propose that a person has a stake in a binary decision if and only if her well-being level for one of the options lies below the threshold, and for the other one above the threshold. All these persons' stakes can be treated as equal (since all that matters is getting as many of them as possible above the threshold — it does not matter how

much they might gain from this). This means that everyone who will be above the threshold with either option does not have a stake, in this sense, in the decision. It also means that those below the threshold with either option have no stakes. That is, we define the criterion of the common good and an individual's stakes as follows.

***Headcount-Sufficientarian Stakes:*** An individual has one stake if and only if her well-being level for one of the options lies below the sufficientarian threshold, and for the other above the threshold.

***The Headcount-Sufficientarian Criterion:*** The option with the greatest number of individuals above a given threshold of well-being is collectively optimal.

Then, by assigning votes in proportion to stakes, such that all stake-holders receive an equal vote, it will be ensured that the option that is selected is the one that lifts a maximum number of people (namely, the majority) above the given threshold. This is the collectively optimal option, according to the headcount-sufficientarian criterion. Thus, by inserting the above two definitions in the original argument from collective optimality (replacing Welfare Stakes and the Sum-Total Criterion respectively), we can easily turn it into *headcount-sufficientarian argument from collective optimality*.

On the other hand, according to the Negative thesis sufficientarian criterion, distributions of individual welfare are ranked by considering exclusively the distribution of those levels of well-being, for the given pair of options, which are below the given threshold. Then, we can define a person's stakes as the difference between her well-being levels, if *both* lie below the threshold — or, if only *one* of these levels lies below the threshold, as the difference between the lower level and the threshold itself. This means that those who will be above the threshold with either option do not have any stakes in the decision. And those who are above it with one option will not have their stakes increased by possible gains above the threshold. We can then define the following:

***Negative Thesis-Sufficientarian Stakes:*** An individual's stakes consist in the differential between her well-being levels for the two options, if *both* levels lie below the sufficientarian threshold, or as the differential between the lower level and the threshold itself, if only *one* of these levels lies below the threshold; her stakes are zero if her well-being levels from either option are at or above the threshold.

***The Negative Thesis-Sufficientarian Criterion:*** The option with the greatest sum-total of individual well-being levels that are below a given threshold is collectively optimal.

Assigning voting weights in proportion to Negative Thesis-Sufficientarian Stakes will ensure that the option that is selected by the weighted majority rule is the one that maximises the sum-total of well-being of all those below the threshold. Then, we can insert Negative Thesis-Sufficientarian Stakes and the Negative Thesis-Sufficientarian Criterion into the original argument and derive a *negative thesis-sufficientarian argument from collective optimality*.

Within this negative thesis-sufficientarian framework, we might propose even more sophisticated definitions of a person's stakes, e.g. as the difference between her  $f$ -transformed well-being levels, if *both* lie below the threshold — or between the lower  $f$ -transformed level and the threshold itself, if only *one* of these levels lies below the threshold. By employing a prioritarian function  $f$  for this transformation, we can ensure that the weighted majority rule satisfies a sufficientarian criterion with prioritarian concerns below the threshold.

So, given the appropriate definition of 'stakes', the weighted majority rule can be shown to be collectively optimal in the sense that it satisfies a number of *sufficientarian* criteria. The same holds for still other criteria. In order to avoid tedious repetition, I now indicate how the *original argument from collective optimality* can be adapted to these.

For instance, a *maximin* criterion ranks distributions of individual well-being according to the levels of well-being of the worst off individual(s) in each distribution. We can compare the welfare level of the worst off individual(s) under option  $x$  with the welfare level of the worst off individual(s) under option  $y$  and propose that only the one(s) with the lower level among them have any stakes. Then, only the worst off individual(s) (across both options) have any votes, and thus, assuming self-interest voting, the weighted majority rule will select whatever option is best for the worst off individual(s). If the worst off individual(s) are equally badly off under both options, one could propose that only the one(s) with the next worst level of well-being have any stakes (or, if they are the same, with the next-to next worst level, and so on). This would render the weighted majority rule collectively optimal in a *leximin* sense of the word.

Moreover, one could render the argument from collective optimality in accordance with non-welfarist criteria.<sup>10</sup> For instance, an *autonomy-constrained* criterion ranks the distributions of well-being according to, e.g., their sum-totals, under the constraint that the outcome may not violate anyone's autonomy. Incorporating such a constraint into the definition of stakes can make the weighted majority rule collectively optimal in an autonomy-sensitive way. One simple suggestion would be to define stakes as proportional to well-being differentials, but add that an individual  $i$ 's stakes can be

---

<sup>10</sup> Such an incorporation of 'non-welfarist considerations' is mentioned, but not further developed, by Fleurbaey (*mimeo*: 6).

silenced or overridden by another individual  $j$ 's, if  $j$ 's autonomy is threatened, were  $i$  to use a stake-proportional number of votes.

To consider an example, let us return to Nozick's case of the *bachelorette* considering which one of her suitors she should marry. This decision undoubtedly affects both her and all her suitors' well-being (and probably a lot of others' as well), so they all have stakes. Should they all get to vote in this decision? This seems counterintuitive, or so Nozick claims.<sup>11</sup>

To get a simple and clear-cut binary decision from this example, imagine that the bachelorette ponders this question: 'Should I marry one of my suitors, Adam, or not?', and that these two options affect the well-being of only these two. If Adam had sufficiently large stakes, such that they outnumber the bachelorette's, he could outvote the bachelorette and get to marry her against her will. But, one might plausibly claim, being forced to marry against one's will violates a person's autonomy — while being forced to *refrain* from marrying someone does not. Since the bachelorette's autonomy thus is in danger, while Adam's is not, on an autonomy-sensitive notion of stakes his stakes are overridden or silenced, so only she gets a vote (or, alternatively, a veto right). The weighted majority rule selects the option containing a smaller sum-total of well-being, but this is in accordance with the proposed autonomy-constrained criterion.<sup>12</sup>

Similar arguments can be made for the collective optimality of the weighted majority rule, for non-welfarist criteria that are sensitive to other alleged values, such as liberty, basic rights or desert.

### 3.3 The generic argument from collective optimality

I have now reconstructed and clarified versions of the *original argument from collective optimality* for a sum-total and a prioritarian criterion of collective optimality (according to the proposals by Fleurbaey and Brighouse and Fleurbaey, respectively). I have moreover specified how to further adapt the argument to sufficientarian, maximin (or leximin), and non-welfarist criteria. The adaptation of the argument to the stated criteria was achieved by redefining the notion of stakes in accordance with the respective criterion.

---

<sup>11</sup> Nozick (1974: 269). Cf. 2.3.2 above.

<sup>12</sup> For a similar idea, concerning the all-affected principle and a rights-constrained notion of 'affected interest', see Arrhenius (2013).

There are several complications with my above suggestion, such as that there may be cases where everyone's autonomy is threatened by everyone else using their stake-proportional votes. Does this mean that autonomy then can be disregarded, or that the degree of autonomy violation, according to some measure, should be minimised — or that that the group should solve the problem by other means than voting? I do not delve into lengthy investigations of these questions here, answers to which must ultimately be derived from the autonomy-constrained criterion of collective optimality.

I now want to state an argument from collective optimality without specifying which version of these premises is employed. That is, I propose what I call the *generic argument from collective optimality*. It shows that the weighted majority rule is collectively optimal for any one of the proposed criteria of the common good. It relies on the premise that stakes are defined appropriately, given a specific criterion (according to my above proposals).<sup>13</sup> We can thus define:

**Generic Stakes:** An individual's stakes are defined in accordance with the given criterion of the common good.

We can then state the *generic argument* as follows.

(1) *The Generic Theorem.* For all individuals and any decision with two options, and for a given criterion of collective optimality, given the Weighted Majority Rule, Generic Stakes, and Self-Interested Voting, the collectively optimal option, according to this criterion, is selected as the outcome.

(2) A rule that selects the collectively optimal option, among all the given alternatives, is collectively optimal.

Hence:

(3) For all individuals and any decision with two options, given Generic Stakes and Self-Interested Voting, the Weighted Majority Rule is collectively optimal, according to this criterion.

This is, no doubt, a promising result: it shows that the weighted majority rule can derive a collectively optimal outcome from purely self-interested input, when the stakes are appropriately defined.

All the above arguments rely on the following assumption.

**Self-Interested Voting:** Every voter (that is, individual who has been assigned a positive number of votes) votes according to her self-interest.

Consider Brighthouse and Fleurbaey's explicit assumption 'that every individual always votes according to his interests'.<sup>14</sup> Self-Interested Voting is slightly weaker, since it concerns not every individual, but only voters. Voters, in turn, are defined as those and only those individuals who have been assigned a positive number of votes. Self-Interested Voting is thus equivalent with Fleurbaey's assumption 'that each voter votes for the option which is more

---

<sup>13</sup> The *generic argument from collective optimality* may cover even other criteria and appropriately defined stakes. Since I have not shown this, I leave this possibility open for further research.

<sup>14</sup> Brighthouse & Fleurbaey (2010: 143–144).

favorable to him’ — if ‘more favorable to him’ is interpreted as ‘better for him’ or ‘in his self-interest’.<sup>15</sup>

Self-Interested Voting is still quite a strong assumption. In the subsequent three chapters, I propose how it can be relaxed. Prior to that, I clarify some points about the *generic argument from collective optimality* and the arguments to come.

### 3.4 Further clarifications

**Externalities.** The conclusion of the *generic argument from collective optimality* states that the weighted majority rule is collectively optimal, according to the given criterion, if the stakes are defined appropriately and voters vote according to their self-interest in binary decisions. Now consider a case in which the given definition of stakes gives at hand that there are stakeholders — who should be assigned an appropriate number of votes, according to the Weighted Majority Rule, and hence are voters, according to the definition of this term — who are incapable of voting. These may be future people, the comatose, very young children or animals, to take just a few rather obvious examples. Then, the decision at hand affects the well-being of those who cannot take part in making it. In other words, there are (positive or negative) externalities.

What do such cases mean for the argument from collective optimality? They do not falsify the theorem but instead violate one of its conditions. Recall that the Self-Interested Voting assumption states that every voter (individual who has been assigned a positive number of votes) votes according to her self-interest. This assumption implies that every voter votes. If certain voters are incapable of voting, the assumption is violated. So for such cases, the theorem's result is not relevant. Since one of its conditions does not hold, we cannot use the theorem to infer the collective optimality of the weighted majority rule. In this study, however, I disregard the problem of externalities.<sup>16</sup> The problem of externalities brings us to the more general problem of abstentions.

**Voter abstention.** As just stated, the result of the *generic argument from collective optimality* is conditional on the Self-Interested Voting assumption, which implies that every voter votes. In other words, voters are assumed not to abstain. In cases where some do abstain, again, the above arguments are

---

<sup>15</sup> Fleurbaey (*mimeo*: 6).

<sup>16</sup> This problem might be solved by assigning representatives to incapacitated stakeholders, as e.g. proposed by Arrhenius (2010). Such a solution needs to be developed in detail, a task that is beyond the scope of this study.

not relevant and the theorem cannot be used to infer the collective optimality of the weighted majority rule. (I briefly return to this issue in the Appendix.)

**Agenda-setting.** The *generic argument from collective optimality* states that the weighted majority rule is collectively optimal, under the given assumptions, since it selects the collectively optimal option among the two options of a given decision. The latter clause does not preclude that there might be yet another option that, in fact, has an even greater sum-total of well-being. Consider a case in which two options are up for decision, while a third option, with, e.g. a greater sum-total of well-being, for some reason is ignored. It might be that there is a small but influential group of people who are made worse off by the third option and who manage to convince everyone else that it is not feasible. So this option does not appear on the decision agenda. Then, obviously, though the weighted majority rule will select the better option of the two on the agenda (if there is one), it will not select the collectively best option.

What does such a case mean for the argument from collective optimality? The argument claims that, for any decision with a given number of options (in this case, two), the weighted majority rule is collectively optimal since it selects the better option among the given ones. Introducing additional options thus falls outside the domain of the argument. (The same holds in Chapter 7, where I consider decisions with more than two given numbers of options: if the collectively best option is not on the agenda, the weighted majority rule will obviously fail to select it.) Hence, I subsequently disregard what we may call ‘incomplete agenda’ cases, where not all available options are up for decision. Moreover, I assume that the agenda is fixed — that the options are given — rather than that they are up for yet another collective decision, to avoid potential regress problems.<sup>17</sup>

**Procedural optimality.** The generic argument from collective optimality is concerned with the optimality of outcomes, that is, of the collectively selected options. But collective optimality may also be affected by the way such decisions are made. Deciding by weighted majority rule may be collectively worse than deciding by means of some other procedure — even though the

---

<sup>17</sup> For a discussion of how the agenda problem affects Brighouse and Fleurbaey's principle of proportionality, see Petré (2007: 112f.). There is another kind of agenda problem often considered in the literature (see e.g. Nurmi 2010). This problem is only relevant for decisions with more than two options. In such cases, the outcome of such a decision may in part depend on the order in which pairs of options face each other. Then, whoever is in control of the serial agenda, which specifies this order, may change the outcome in her favour. This problem of ‘serial agenda control’ is here not of relevance, since I for now consider only decisions with two options, where there is only one possible pair of options to face each other. I briefly return to this problem in Chapter 7, though, where I consider decisions with more than two options.

former selects better outcomes. This may be, e.g., because the former is more costly (requiring substantial resources to assess stakes and distribute votes) or less publicly accepted (voters are suspicious of others receiving too many votes) or what have you. Eventually, such ‘side-effects’ must be taken into account when making an argument for the *overall* collective optimality of the weighted majority rule. In the present study, I largely disregard such procedural gains and losses and only claim to argue for the collective optimality of this rule when looking at its selected outcomes.

***If we know all the stakes, why vote?*** Assigning the right number of votes to people presupposes correct stake assessments. In order to implement the weighted majority rule, then, some *social planner* would need to assess their stakes for the given options. This could be an individual or an institution or a sophisticated computational device. Either way, assigning votes presupposes the measurement, as well as the interpersonal comparison, of individual well-being. It might appear that the kind of measurements and comparisons that are needed for an assessment of the individual stakes are the same as the ones required for directly applying the relevant criterion of collective optimality to compare the options. So one might wonder why the social planner should bother to assign votes rather than go straight for implementing the collectively best option. If all the stakes are known to the social planner, why bother taking a vote — not least when voting is costly?

However, as it turns out, appearances are mistaken on this point. Less information is required to assess individual stakes than to assess which option is collectively best. Consider for instance the *sum-total* criterion. In 2.3.2 above, I stated that, in order to make the claim that ‘Option *x* is collectively better than option *y*’, this criterion presupposes that we can for each individual, (1) rank the options according to her well-being levels, and (2) count the units of her well-being differential between the options; that we can, within the group, (3) meaningfully compare these differentials among all individuals; and for each option, (4) add the units by which it is higher-ranked than each of its alternatives by any of the group's individuals. Then, the options can be collectively ranked in accordance with the number of units we added up for each of them. Now, in order to assess the individuals' stakes, however, we do not need to presuppose all four claims. All we need to measure and interpersonally compare is individual well-being differentials for the given pair of options. We do not need to know the individual rankings of these options. In fact, claims (2) and (3) suffice. The information required in claim (1) is provided by the voter herself — when she, as assumed, votes accord-

ing to her self-interest.<sup>18</sup> And the aggregation referred to in claim (4) is done by the mechanism of the weighted majority rule.

The same holds even when stakes are defined in the light of a *non-welfarist* (e.g. autonomy-constrained) criterion: what is required is information about each individual's interpersonally comparable well-being differentials for the properly constrained options — not information about which option is better for each. So again, claims (2) and (3) suffice.

Again, the same holds for the *prioritarian* criterion, which (as stated in 2.3.2 above) in addition to claims (1) through (4) presupposes that well-being differentials can be ordered along a prioritarian dimension for discounting the better off individuals' stakes. Assessing the voters' prioritarian stakes presupposes only claims (2) and (3), along with proper discounting. As Brighouse and Fleurbaey explicitly note, this assessment does *not* presuppose knowledge of which option is the better one for each individual, as some might claim.

‘This claim is incorrect because the stakes measure the intensity but not the direction of individual interests. In many contexts [...] the relative size of stakes is roughly known but not the preference of the individual, so that a vote is needed to reveal the latter’.<sup>19</sup>

A similar line of reasoning applies for the *negative-thesis sufficientarian* criterion. Recall that an individual's stakes were defined as the difference between her well-being levels for the pair of options if *both* levels lie below the threshold of sufficiency — or between the lower level and the threshold itself if only *one* of these levels lies below the threshold. So assessing the voters' stakes presupposes claims (2) and (3), along with an identification of the threshold above which well-being differentials should be disregarded. It is not necessary to know the better option for each individual. Again, less information is required to assess the stakes than to apply the criterion of the common good directly to the options.

This advantage of using the weighted majority rule is also retained when stakes are defined in accordance with a *headcount sufficientarian* criterion. Recall (again from 2.3.2 above) that it was stated to presuppose a threshold such that we can 1) identify and count all individuals who are below the threshold-level of well-being with one option and above this level with the other and 2) compare the number of individuals who are above this threshold with one option with the number who are above the level with the other option, to rank the options in accordance with these numbers. When assessing the stakes, we only need to identify, as holders of equal stakes, all individu-

---

<sup>18</sup> In the next three chapters I show that this assumption can be relaxed. This then shows that using the weighted majority rule is still less demanding than straight application of the collective-optimality criterion.

<sup>19</sup> Brighouse and Fleurbaey (2010: 146).

als who are below the threshold-level of well-being with one option and above this level with the other. All the remaining necessary information is then disclosed by the operation of the weighted majority rule, under the assumption of self-interested voting.

However, note that this informational advantage might not be retained for the maximin and leximin criteria, as stated. For *maximin*, we need to identify, among all the given options, the individual who is worst off. If there is only one individual on the lowest level, she holds the only stakes in the decision and we need no information about how she ranks the options (this information is instead disclosed by her vote). Yet if there is more than one individual on the lowest level, we need to know whether they are made worst off by the same option or by different options. In the latter case, we would have to conclude that no one holds any stakes (considering that the options are ranked as equally good according to maximin). Then, no one gets any votes and the weighted majority rule cannot select any option. (However, as both options are ranked as equally good, any one of them may then be selected by some secondary decision rule.)<sup>20</sup>

This observation carries over to the *leximin* criterion, which differs from maximin only when several individuals are made equally worst off by different options. And this, again, we can ascertain only when given information about the individual rankings.

Still, from these considerations it should now be clear that knowing all the stakes does not necessarily turn voting into a redundant exercise.

**Stake assessment.** As just noted, the implementation of the weighted majority rule presupposes a *social planner* who assesses the individual stakes for the given options. In order to know who has stakes and how great they are, the planner would need to know which theories of well-being and the common good are relevant or correct. But this just delegates the main problem of applying the weighted majority rule to the undefined entity of a ‘social planner’. How should this planner operate? Is there not going to be general disagreement about which theories of well-being and the common good are relevant or correct? Should we then take a collective decision about the proper definition of stakes? And how should the stakes in *this* decision be determined, prior to the collective decision on how to define stakes? Do we face a regress, or simply a dead end?

---

<sup>20</sup> Note though that it might be proposed that all the equally worst off should be identified as stake-holders, no matter under which option they are made worst off. This would serve a slightly different maximin criterion, which includes a secondary principle of minimising the number of the worst off when either option makes some individuals equally worst off. Such an *extended maximin* criterion would not require information about individual rankings of the options.

Defining the stakes is a difficult theoretical issue, unlikely to be settled by conclusive arguments or general agreement anytime soon. Moreover, even if we somehow could settle for the relevant or correct theories of well-being and the common good, there still are the practical problems of the measurement and interpersonal comparability of the stakes.

However, we should not forget that *any* proposed distribution of votes, and *any* justification — in terms of stakes or otherwise — are sources of potential disagreement. As, e.g., the Stockholm congestion tax referendum (see 1.1 above) clearly showed, the alleged equal vote is not equal: there are always some who are without a vote. The drawing of any on-off boundary of vote assignment is potentially contested — as is the equal weighting of the votes of those included within the demos. (This was the problem of the tyrannical majority case, in which an almost unaffected majority was able to dominate a highly affected minority, see 1.1 above.) The problems are basically the same, regardless of whether the vote is allegedly equal or unequal. Yet in contrast to an apparently simple ‘one person one vote’ principle, the weighted majority rule forces us to acknowledge these problems — and moreover provides us with a principled way of addressing them.

If there is overlapping consensus about what makes people well off, pragmatic agreement on who has stakes in a given decision might not be far off.<sup>21</sup> How great would the respective stakes be? As already noted, the application of the weighted majority rule inherits some of the measurement and interpersonal comparability problems faced by the application of the relevant criterion of the common good (though it does not inherit all of the problems, as it does not presuppose information about the individual rankings of the options in terms of self-interest). Moreover, depending on which theory of well-being is presupposed, estimating and comparing people's stakes will be more or less difficult. Consider an objective list theory: if it lists items that can be operationalised, stake assessment might be quite precise. Subjectivist theories such as desire fulfilment or hedonism pose the problem of objective assessment of mental states, such as experiencing pleasure or desiring something (in the well-being relevant way). Still, there are ways — such as asking people, observing their behaviour, scanning their brains — to deal with the problem.

However, in real life there will be uncertainty about the exact stakes, no matter which theories are presupposed and no matter how much resources we invest into measuring and comparing (especially since such resource investment must, eventually, be taken into an overall account of the collective optimality of deciding by weighted majority rule). On this note, I want to briefly refer to Fleurbaey's proposal that, when there is uncertainty about

---

<sup>21</sup> Cf. the experimental results by Dimdins et al. (2011), suggesting that people are prepared to accept weighted votes when they perceive the voters to be unequally affected (as referred to in 1.2 above).

the exact individual stakes, then the weighted majority rule can be shown to be *expectedly* collectively optimal, if the voting weights are proportional to *expected* individual stakes.<sup>22</sup> This gives us the outline of yet another argument for the weighted majority rule, which we may call the *argument from expected collective optimality*. This argument is epistemically less demanding of the ‘social planner’, and hence potentially more relevant for real-life decision-making. (This comes at the price of a somewhat weaker conclusion in terms of *expected* collective optimality.)

In praxis, stakes might be estimated by means of some proxy, such as ‘location, occupation, or financial situation’.<sup>23</sup> Estimations of the stakes could alternatively be obtained from the individual voters themselves. For instance, people could be allowed to trade other resources to attain any number of votes that corresponds to their stakes. That might involve buying votes with money<sup>24</sup> or being allowed to fill in as many ballots as they wish, where ballots are deliberately designed to be effort- or time-consuming. This presupposes that each voter can appropriately estimate not only the ‘direction’ of her stakes, but also their relative size, or at least make only stake-appropriate trades. Moreover, it presupposes that people value equally the required resources (such as money or time).<sup>25</sup>

One further note: some people object to employing the weighted majority rule on the grounds that it would open the door to all kinds of manipulation, as voters may try to gain advantages by misrepresenting their stakes. Still, nothing of what has been said so far presupposes that the voters themselves get a (final) say in determining their stakes. Certainly there are ways of determining stakes that do not rely on the voters' own assessments alone. For instance, there are general and systematic results from the social sciences, concerning the effects that different kinds of outcomes typically have on people. Such research can — and frequently does — inform us how specific groups of people, such as the poor, the very wealthy, the unemployed or small business owners, are affected by certain options (e.g. in terms of well-being). Thus, it seems that there are ways to prevent or expose and counteract stake manipulation.

In the remainder of this study, however, I disregard the problems of stake assessment. My focus is the performance of the weighted majority rule, given the correct stake assessment.

---

<sup>22</sup> Fleurbaey (*mimeo*: 16).

<sup>23</sup> Fleurbaey (*mimeo*: 8).

<sup>24</sup> However, see Downs (1957) for some cautionary remarks on the vote market.

<sup>25</sup> Alternatively, people could simply be allowed to trade votes on different decisions — this would come close to Buchanan and Tullock's (2004) idea of the collectively beneficial effects of logrolling (cf. 7.1 below).

### 3.5 Conclusions

In this chapter, I have reconstructed the *original argument from collective optimality* proposed by Fleurbaey and Brighthouse and Fleurbaey, respectively. They make a case for the weighted majority rule from a sum-total and a prioritarian criterion of the common good, respectively. I have suggested that their theorems can be reconstructed along the lines of a formal model, which provides insight into the operation and output of this rule, given certain assumptions. Moreover, I have argued that their arguments, resting on these theorems, can be generalised for a larger variety of criteria of the common good. I have proposed the following *generic argument from collective optimality*.

(1) *The Generic Theorem*. For all individuals and any decision with two options, and for a given criterion of collective optimality, given the Weighted Majority Rule, Generic Stakes, and Self-Interested Voting, the collectively optimal option, according to this criterion, is selected as the outcome.

(2) A rule that selects the collectively optimal option, among all the given alternatives, is collectively optimal.

Hence:

(3) For all individuals and any decision with two options, given Generic Stakes and Self-Interested Voting, the Weighted Majority Rule is collectively optimal, according to this criterion.

The Generic Stakes assumption has been considered in some detail in the present chapter. I have provided a variety of specifications of ‘stakes’ and thereby adapted the argument to a variety of criteria of the common good. The Self-Interested Voting assumption is more closely examined and relaxed in the following three chapters.

Finally, I have stated a number of limitations of my study, namely, that it disregards the problems of externalities, of voter abstention, agenda-setting, and procedural costs and benefits of implementing the weighted majority rule, as well as problems concerning the how to assess individual stakes in practice. I have, moreover, argued that voting is not necessarily made redundant by the knowledge of the stakes presupposed by the weighted majority rule.

## 4 Self- and common-interested voting

### 4.1 Introduction

As the previous chapter has shown, the *generic argument from collective optimality* for the weighted majority rule rests on the following assumption:

***Self-Interested Voting:*** Every voter votes according to her self-interest.

As noted, this is a rather strong assumption. This chapter's aim is to construct an argument to the effect that the assumption can be relaxed (rendered logically weaker), while retaining the collective optimality of the weighted majority rule. To be precise, in 4.2 I show that this rule is collectively optimal even when all of the voters vote according to the common interest, and even when there is a mix of self- and common-interest voters. I then address a claim to the contrary, which is known as the 'mixed motivation problem'. I show that the alleged problem is misconceived and that it does not constitute an objection to the here proposed argument (in 4.2.1). In 4.3, I examine the conditions under which the weighted majority rule is at risk of failing to select the collectively best option. I moreover suggest a motivational background of the relaxed Self-Interested Voting assumption, in terms of the voters' motivating desires and relevant beliefs (in 4.3.1). Section 4.4 concludes.

### 4.2 The extended argument from collective optimality

Imagine a given group of voters who face a binary issue with options  $x$  and  $y$ . We can then go step by step through a number of scenarios in which we vary the voters' voting behaviour.<sup>1</sup> (Note that, just like the *generic argument*

---

<sup>1</sup> I owe the outline of this argument to Marc Fleurbaey (personal communication, May 2009). The argument is also hinted at in Brighthouse and Fleurbaey (2010: 152): '[Common-interested voting] need not be problematic because, on the whole, it will reinforce the weighted majority in favor of the good options'.

from *collective optimality*, the following argument can be stated for any of the criteria of the common good I have discussed in 2.3.2 above.)

**Only self-interest voters.** In this scenario, each voter casts her vote(s) for the option that is in her self-interest. Moreover, the stakes are defined in accordance with the relevant criterion of the common good (this assumption is taken to be satisfied throughout the stated scenarios in this chapter). Then, as shown by the *generic argument from collective optimality*, the weighted majority rule will select the common-interest option — let us say this is  $x$  — and thus be collectively optimal.

**One common-interest voter.** Now, we change the above scenario just a bit, such that all voters except one vote in their self-interest. The one voter,  $i_1$ , instead casts her stake-proportional number of votes for the alternative that is in the common interest. As assumed above, this is  $x$ . Then, there are two possibilities: either (1)  $i_1$  votes exactly as she does in the *Only self-interest voters* case. This means that she votes for  $x$  in both cases —  $x$  is both in her self-interest and in the common interest. This will not change the total numbers of votes for  $x$  and  $y$ , respectively. So the outcome by weighted majority rule is still  $x$ . Or (2)  $i_1$  now votes differently from the *Only self-interest voters* case. This means that she votes for  $y$  in that case. Changing her vote to  $x$  will not reduce the total number of votes for the common-interest alternative. Instead it will increase it by the amount of  $i_1$ 's vote(s). Thus, the outcome by weighted majority rule again is  $x$ , the common-interest option.

**Two common-interest voters.** Let us make another little change to the scenario, such that all voters except two vote in their self-interest. These two voters,  $i_1$  and  $i_2$ , cast their votes for the common-interest option  $x$ . For each of them there are again two possibilities: either (1)  $i_1$  (or  $i_2$ ) votes as she does in the *Only self-interest voters* case. This means that they vote for  $x$  in both cases —  $x$  is both in her self-interest and in the common interest. Again, this will not change the total numbers of votes for  $x$  and  $y$ , respectively. So the outcome by weighted majority rule is still  $x$ . Or (2)  $i_1$  (or  $i_2$ ) now votes differently from the *Only self-interest voters* case. This means that they vote for  $y$  in that case. Changing their votes to  $x$  will not reduce the total number of votes for the common interest alternative. Instead it will increase it by the amount of  $i_1$ 's (or  $i_2$ 's) vote(s). Thus, the outcome by weighted majority rule again is  $x$ , the common-interest option.

We can now proceed to *Three common-interest voters*, *Four...*, *Five...*, and so on, all the way to *Only common-interest voters*, where all voters cast their votes in the common interest. Every time, the same kind of reasoning applies. For each case, it is shown that the weighted majority rule still selects  $x$ , the common-interest option. This argument does not rest on the strong Self-Interested Voting assumption but rather on the following weaker claim.

**Self- Or Common-Interested Voting:** Every voter votes according to her self-interest or according to the common interest.

The argument then shows that the following Extended Theorem is true, which in turn can be employed in what we may call the *extended argument from collective optimality*.

(1) *The Extended Theorem.* For all individuals and any decision with two options, and for a given criterion of collective optimality, given the Weighted Majority Rule, Generic Stakes, and Self- Or Common-Interested Interested Voting, the collectively optimal option, according to this criterion, is selected as the outcome.

(2) A rule that selects the collectively optimal option, among all the given alternatives, is collectively optimal (*ceteris paribus*).

Hence:

(3) For all individuals and any decision with two options, given Generic Stakes and Self- Or Common-Interested Interested Voting, the Weighted Majority Rule is collectively optimal (*ceteris paribus*), according to this criterion.

Can we relax the Self- or Common-Interested Voting assumption further, without jeopardising the collective optimality of the weighted majority rule? For instance, could ‘partial-interest’ voting behaviour be allowed? This is displayed by a voter who casts her vote(s) in the collective interest of a subgroup of the entire group (that is, a proper subset of all voters, which is non-identical to the subset containing only herself, but which may — or may not — include herself). It can easily be seen from the following scenarios that this can falsify the theorem, in the sense that, given this relaxed assumption, the result of collective optimality might not hold.

**One partial-interest voter.** Let us assume that all voters except one cast their votes in either their self- or the common interest and that  $x$  is the common interest. The weighted majority rule would pick  $x$  if everyone was voting either in their self- or the common interest. However, the excepted voter,  $i_3$ , is a partial-interest voter: she casts her vote in the collective interest of her family. Moreover, assume that  $i_3$  is pivotal in this decision. A voter is *pivotal* in a binary decision if and only if, for a given distribution of all the voters' votes between the two options, had this voter voted for the other option, the outcome would have changed to that option. While  $x$  is in  $i_3$ 's self-interest,  $y$  is in  $i_3$ 's family's collective interest. So since she votes partial-interested in the way described,  $i_3$  votes for  $y$ . And since she is pivotal, the weighted majority rule will then pick  $y$ , which is not in the common interest.

The same holds for ‘selfless’ voting behaviour that is displayed by voters who vote in everyone's collective interest but their own. Selfless voting is really just a form of partial-interest voting. The following scenario shows that selfless voting behaviour may falsify the above theorem as well.

**One selfless voter.** Assume that all voters except one cast their votes either in their self-interest or in the common interest. The excepted voter,  $i_d$ , has enormous stakes. Everyone else's stakes are rather small though, such that the sum of their stakes is outweighed by  $i_d$ 's stakes. Then, the weighted majority rule assigns a large number of votes to  $i_d$ , which outnumber everyone else's votes. This implies that  $i_d$  is pivotal. Now, assume that  $i_d$ 's self-interest is  $x$ , while everyone else's collective interest is  $y$  and that  $i_d$  is a selfless voter. Then,  $i_d$  votes for  $y$ . And since she is pivotal, the weighted majority rule will select  $y$ , which is not in the common interest.

The same also goes for voters who cast their votes according to other ends — such as for the promotion of aesthetic value, or according to their duties, or in accordance with the wrong theory of welfare — whenever these ends conflict with either these voters' self-interest or the common interest, properly conceived. The weighted majority rule's collective optimality — the property of selecting the common-interest option — has only been shown conditional on self- or common-interested voting.

#### 4.2.1 Rebutting the 'mixed motivation' problem

Now, I have already noted that in cases where all voters have equal stakes (*equal-stakes cases*), the weighted majority rule is extensionally equivalent with the *simple* majority rule. Concerning this latter rule, the above conclusion concerning collective optimality has been contested by Jonathan Wolff, who states that 'it can easily be demonstrated that, if part of the electorate vote in pursuit of their own interests, and part for common good [defined as majority interest], then it is possible to arrive at a majority decision which is neither in the majority interest, nor believed by the majority to be for the common good'.<sup>2</sup> Wolff establishes this conclusion with an example, which I restate as follows.

**Mixed motivation.** There is a group facing two options,  $x$  and  $y$ .  $x$  is in the self-interest of 40% of the group (I call this the first subgroup), while  $y$  is in the self-interest of 60% (the second subgroup). Moreover, 80% of the entire group (across the first and second subgroup) believe  $y$  to be the common good, defined as majority interest, while 20% believe this of  $x$ . Now, suppose that all voters within the first subgroup vote according to their self-interest, while all the voters within the second vote according to their perception of the common interest. Then, 52% will vote for  $x$  — the option that is *not* in the majority interest and *not* believed to be so by a majority of the group. (This can be easily seen, as all in the first subgroup — 40% of the group — will vote for  $x$ , as well as the 20% of those within the second subgroup — constituting 60% of the group — who take  $x$  to be the common good, which amounts to an additional 12%.)

---

<sup>2</sup> Wolff (1994: 194).

While the example is certainly correct and seems to support Wolff's conclusion, the conclusion is misleading: it obfuscates the real problem. As we can see from the description of the example, there are *two* assumptions at play: voters have mixed motivation *and* 80% of the group are correct in their assessment of the common good — while 20% are incorrect. So Wolff's conclusion should rather state: 'if part of the electorate vote in pursuit of their own interests, and part for common good [defined as majority interest]', *and if some of them are mistaken in their beliefs about the common good*, 'then it is possible to arrive at a majority decision which is neither in the majority interest, nor believed by the majority to be for the common good'.

My above *extended argument from collective optimality* shows that the weighted majority rule selects the common-interest option — or in Wolff's terms, the majority interest. In equal-stakes cases, the weighted majority rule is extensionally equivalent to the simple majority rule, thus the same result holds for the latter rule. What the argument assumes is mixed motivation (and correct stake-assignment). This shows that the mixed motivation assumption does not do the job Wolff claims it to do.

In fact, the mixed motivation assumption is not relevant at all within Wolff's conclusion. To see this, consider the following case.

***Homogeneous motivation.*** There is a group facing two options,  $x$  and  $y$ .  $x$  is in the self-interest of 40% of the group (the first subgroup), while  $y$  is in the self-interest of 60% (the second subgroup). Now suppose that all are motivated by their self-interest. Within the first subgroup, all correctly perceive  $x$  to be in their self-interest. Yet within the second subgroup, 20% mistakenly judge  $x$  to be in their self-interest. All vote accordingly. Then, 52% will vote for  $x$  — the option that is *not* in the majority interest. (This can be easily seen, as all within the first subgroup — 40% of the group — will vote for  $x$ , as well as the 20% of those within the second subgroup — constituting 60% of the group — who mistakenly judge  $x$  to be in their self-interest, which amounts to an additional 12%.)

This scenario brings out that it is the assumption of *epistemic failure* that drives Wolff's conclusion that 'it is possible to arrive at a majority decision which is [not] in the majority interest'.<sup>3</sup> Mixed motivation has nothing to do with Wolff's problem. (I return to Wolff's alleged mixed motivation problem when discussing competence assumptions in 5.2.5 below.)

---

<sup>3</sup> Wolff (1994: 194). (Note that we could easily add the assumption that 80% of the entire group (across minority and majority stake-holders) believe  $y$  to be the common good to also derive the second part of Wolff's conclusion that the simple majority rule selects an option *not* 'believed by the majority to be for the common good.')

For a similar analysis of Wolff's problem, see Graham (1996).

### 4.3 Erratic voting behaviour

I now want to examine the conditions under which the weighted majority rule's collective optimality is not guaranteed by the *extended argument from collective optimality*. For the sake of simplicity, I presuppose a sum-total criterion of the common good.

Recall that we are here considering decisions with two options,  $x$  and  $y$ , and thus there are only two sets of voters. There is the set of those who are better off with  $x$  and the set of those who are better off with  $y$ . (Recall that individuals who are equally well off with either option are not voters as the term is defined.) Moreover, recall that we are especially interested in cases where one of the options is collectively better than the other. (If both are equally good, it does not matter from a common good perspective which option is selected.) In these cases, the sum-total of stakes within one of two sets of voters must be greater than the sum-total of stakes within the other (Recall that 'stakes' as defined trace the information on which 'collectively better' is defined). Let us then call voters of the set that holds the majority of stakes the *majority stake-holders*, and voters of the other set accordingly the *minority stake-holders*. (Note that the *voters* we call majority stake-holders may be outnumbered by the minority stake-holders — it is the former's *stakes* that in sum outnumber the latter's.) It can now be seen that the majority stake-holders' self-interest coincides with the common interest, while the minority stake-holders' self-interest is opposed to the common interest.

Let us call a voter who votes neither in her self-interest nor in the common interest, and thereby violates Self- Or Common-Interested Voting, an *erratic voter* or a voter who *votes erratically*. Now, within the present binary decision context we can see that a minority stake-holder *cannot* be an erratic voter, regardless of how she casts her vote(s) (as long as she does vote, as here assumed). If she does not vote for the option that is in her self-interest, this implies that she votes in the common interest; if she does not vote for the option that is in the common interest, this implies that she votes in her self-interest. Or, to put it differently: If exactly one of the two given options is in the common interest, then, since the minority stake-holders' self-interest is opposed to the common interest, the other option must be in the latter's self-interest. So regardless of how they vote, they always vote either in their self- or in the common interest.

A majority stake-holder, on the other hand, *can* be an erratic voter, violating Self- Or Common-Interested Voting: by voting for the option that is neither in her self-interest nor in the (coinciding) common interest. Still, such a voter jeopardises the weighted majority rule's collective optimality only if she *also* happens to be pivotal when casting her vote(s) — otherwise her vote(s) do not change the outcome. Thus, the above *extended argument from collective optimality* is true even if we replace Self- Or Common-Interested Voting with the following even weaker assumption:

***Pivotal Self- Or Common-Interested Voting:*** Every voter votes according to her self-interest or according to the common interest, *or is non-pivotal*.

As just stated, this assumption is violated if and only if there is (at least) one pivotal *and* erratic voter. And, as just stated, an erratic voter must necessarily be a majority stake-holder, since minority stake-holders cannot be erratic (as long as they do vote). So we can relax this assumption further, by limiting it to majority stake-holders rather than the entire group of all voters.

***Pivotal Non-Erratic Majority Stake-Holders:*** Every pivotal majority stakeholder votes non-erratically, that is, votes according to her self-interest or according to the common interest.

Until now we have focused on the assumption *that* voters vote in certain erratic or non-erratic ways, but we have not said anything about *why* they would vote in a particular way. Can we make sense of the behavioural assumption by providing a more detailed picture of the underlying voter motivation that might explain this assumption? In the remainder of this chapter, I address this latter question.

#### 4.3.1 Why a voter may vote erratically

Why would a voter vote in accordance with the Self- Or Common-Interested Voting assumption and thereby vote non-erratically? Why would she fail to satisfy this assumption and thereby vote erratically?

In his article on the weighted majority rule, Fleurbaey briefly ponders the problem of ‘erratic voters’. Since his version of the argument from collective optimality is built on the assumption that voting behaviour is solely self-interested, he defines ‘erratic voting’ in terms of voting against one’s self-interest. He speculates that erratic voting ‘may be due to a discrepancy between [the voters’ self-interest] and their preferences [...], or to an informational problem that gives them mistaken beliefs about their [self-]interests’.<sup>4</sup>

From this suggestion we can reconstruct two plausible reasons for erratic voting behaviour in my sense (voting against one’s self-interest or against the

---

<sup>4</sup> Fleurbaey (*mimeo*: 19). Note that, as a possible solution to the problem of erratic voting, Fleurbaey suggests a double-weighted majority rule with ‘optimal weights which are jointly proportional to the stakes and to an index of “reliability” of the voter’ (*mimeo*: 20). Such a rule would, for instance, assign a negative stake-proportional voting weight to a notoriously erratic voter who always gets it wrong. Fleurbaey, however, quickly dismisses this suggestion as being of ‘no practical interest’, but merely serving ‘as a theoretical clarification’. His reason for this dismissal is that negative voting weights would ‘make the voting rule highly manipulable’ (*mimeo*: 20). Fleurbaey’s idea is that, realising that she is assigned a negative voting weights, a self-interested voter will be motivated to vote *against* the option that is in her self-interest, that is, to vote erratically. I do not consider Fleurbaey’s suggestion here since it introduces a decision rule which is clearly distinct from the object of this study: the (stake-) weighted majority rule.

common interest). (1) A voter is motivated by something else than her self-interest or the common interest (according to the relevant theories of well-being and the common good), and acts according to her motivation. (2) The voter is motivated by her self-interest or the common-interest (according to these theories), but has incorrect beliefs about which option is in her self-interest or the common interest respectively and acts according to this belief-desire pair.

To Fleurbaey's pair we may add a third plausible reason for erratic voting behaviour. (3) The voter is motivated by her self-interest or the common-interest and has correct beliefs about what constitutes this self- or common interest respectively, but fails to act according to this belief-desire pair for some reason. It may simply be that she makes a mistake when setting about to act as intended, she may will be weak-willed, or she may suffer from a 'black-out' or other such internal impediments. Alternatively, she may face external obstacles, such as deception or coercion.

From these suggestions, we can reconstruct three assumptions about the motivational set-up of the voters. In conjunction, these assumptions imply the Self- Or Common-Interested Voting assumption.

***Self- Or Common-Interested Motivation:*** Every voter is either self-interested, in the sense that she desires to promote her self-interest, or common-interested, in the sense that she desires to promote the common interest.

***Competence:*** Every voter has a correct belief that voting for some option  $x$ , among the given alternatives, promotes her self-interest (if she is self-interested) or the common-interest (if she is common-interested).

***Success:*** Every voter acts according to the stated belief-desire pair.

The Self- Or Common-Interested Motivation assumption refers to the voters' ends. Note that 'self-interest' and 'common interest' should be interpreted in accordance with the relevant theories of well-being and the common good. Thus, a voter who believes she is motivated by her self-interest, but has a wrong conception of what is good for her is in the present terminology not a self-interested voter. Note also that it is sufficient to read this assumption as a *de re* claim — a voter who desires to promote what is *in fact* her self-interest without (*de dicto*) desiring to promote *whatever it is* that is in her self-interest does satisfy Self- Or Common-Interested Motivation.

The Competence assumption refers to the voters' competence concerning their ends. Note that 'has a correct belief' does not have to be interpreted as 'knows', nor as 'has a justified true belief'. Note also that this assumption only excludes *empirical* uncertainty, regarding what option promotes the voter's end, namely her self- or the common interest. As I just stated, *evaluative* uncertainty regarding the correct interpretation of well-being and the

common good is covered by the Self- Or Common-Interested Motivation assumption.

Finally, the Success assumption ensures that the belief-desire pair referred to in the other two claims is relevant and effective in determining voting behaviour. In the remainder of this study, I simply take the Success assumption for granted. I do, however, briefly return to it in my final discussion in Chapter 8.

It may now be objected that my claim that these three assumptions imply Self- Or Common-Interested Voting is wrong. This can be seen from the following line of reasoning. Consider a group of self-interested voters, such that Self- Or Common-Interested Motivation is satisfied. Then, when deciding how to cast their votes, all are motivated by their desire to promote their self-interest. Yet, it may now be pointed out, on the prior decision on whether to vote or abstain, the assumed desire to promote their self-interest would induce a number of them — if not all — to abstain rather than vote.

To be sure, it would be in a voter  $i$ 's self-interest to abstain in a collective decision if  $i$ 's potential benefit from this decision's outcome — that is, her gain in well-being, if her preferred option would win — were smaller than the cost of voting — that is, her loss in well-being from, e.g., going to the polling station or getting informed about the alternatives rather than engaging in alternative activities that would make her better off.<sup>5</sup>

Or imagine that  $i$ 's share of votes were only a tiny fraction of all votes. Then, it would be rather unlikely that her vote(s) — among, say, millions of others — would make a difference to the outcome. This can be stated in terms of the vote's chance of being *pivotal* in a decision. We may say that a *vote* is pivotal in a binary decision if and only if, for a given distribution of all the votes between the two options, had this vote been cast for the other option, the outcome would have changed to that option. The chance of one particular vote being pivotal decreases rapidly, with increasing numbers of votes. (If the vote were part of a relatively large vote bundle that, because of its size, has a great chance of being pivotal, we may say that the vote inherits the entire bundle's chance of being pivotal: if this vote were cast for other the option, this would mean that all the other votes within the bundle were cast for the other option, and hence there would be a great chance of changing the outcome to this other option. Yet if this vote, as assumed, is a tiny fraction of a total of millions of votes, its chance of being pivotal must be very small.)

Now, if we add a New Competence assumption — stating that voter  $i$  has a correct belief that voting is costly and that the benefits of the outcome are smaller than the costs, or that her vote(s) are not pivotal (or that she is quite likely to have such a belief) — and the Success assumption that the voter

---

<sup>5</sup> On the costs of voting, cf. Downs (1957).

acts according to her belief-desire pair, we must conclude that voter  $i$  will abstain (or quite likely abstain).<sup>6</sup>

However, what this objection shows is not that my above three assumptions do not imply Self- Or Common-Interested Voting. It merely shows that that the latter is not implied by Self- Or Common-Interested Motivation, Success, and *New* Competence. Yet it may be claimed that my Competence assumption, that every voter has a correct belief about *which way* of voting promotes her self- or the common interest, implies that every voter has a correct belief about *whether* voting promotes her self-interest or the common interest. Then, my three stated assumptions may indeed imply abstention and thus cannot imply Self- Or Common-Interested Voting. In order to dodge this difficulty, in the following chapters I take my three assumptions — Self- Or Common-Interested Motivation, Competence, and Success — to be limited to the voters' decisions on which option to vote for. In other words, I stick to the assumption that all voters do vote (see 3.4 above). (This means that the voters' decision to vote rather than abstain would have to be grounded in other, presently unspecified assumptions. In the Appendix, I briefly return to this issue.)

## 4.4 Conclusions

The main argument in this chapter shows that the argument from collective optimality does not need to rely on the strong Self-Interested Voting assumption, that all voters vote according to their self-interest. Instead, the following weaker assumption suffices.

***Self- or Common-Interested Voting:*** Every voter votes according to her self-interest or according to the common interest.

Thus, it turns out that ‘mixed motivation’ does not endanger the collective optimality of the weighted majority rule. It has also been shown that this assumption cannot be rendered more permissive than that — it cannot allow partial-interested voting, including selfless voters, or voters with other, conflicting ends. Some additional reflection revealed that erratic voting, that is, voting that violates Self- Or Common-Interested Voting, jeopardises the weighted majority rule's collective optimality only when it is performed by a pivotal voter. Thus, the behavioural assumption can be further weakened.

***Pivotal Self- Or Common-Interested Voting:*** Every voter votes according to her self-interest or according to the common interest, *or is non-pivotal*.

---

<sup>6</sup> On the so-called ‘voter paradox’, that predictions deduced from assumptions such as Self-Interested Voting, Competence and Success do not correspond to observed voting behaviour, see e.g. Downs (1957: 267), and Sen (1990: 34–36). Cf. my Appendix below.

Moreover, it was shown that minority stake-holders cannot be erratic (as long as they do vote), so one can relax the assumption still further by limiting it to majority stake-holders.

***Pivotal Non-Erratic Majority Stake-Holders:*** Every pivotal majority stakeholder votes non-erratically, that is, votes according to her self-interest or according to the common interest.

Finally, I sketched a possible motivational picture behind the Self- Or Common-Interested Voting assumption. I suggested that it is implied by the conjunction of three assumptions concerning the voters' beliefs and desires.

***Self- Or Common-Interested Motivation:*** Every voter is either self-interested, in the sense that she desires to promote her self-interest, or common-interested, in the sense that she desires to promote the common interest.

***Competence:*** Every voter has a correct belief that voting for some option  $x$ , among the given alternatives, promotes her self-interest (if she is self-interested) or the common-interest (if she is common-interested).

***Success:*** Every voter acts according to the stated belief-desire pair.

In the next chapter, I focus on the Competence assumption and show how it can be relaxed. I eventually return to the question of how this affects the Self- Or Common-Interested Voting assumption.



# 5 Less than fully competent voters

## 5.1 Introduction

In the previous chapter, I suggested that the *extended argument from collective optimality* can be made even when voters do not vote according to their self-interest. Instead, it suffices to assume the following.

***Self- Or Common-Interested Voting:*** Every voter votes according to her self-interest or according to the common interest.

I suggested that this assumption is implied by the conjunction of three assumptions concerning the voters' motivating beliefs and desires.

***Self- Or Common-Interested Motivation:*** Every voter is either self-interested, in the sense that she desires to promote her self-interest, or common-interested, in the sense that she desires to promote the common interest.

***Competence:*** Every voter has a correct belief that voting for some option  $x$ , among the given alternatives, promotes her self-interest (if she is self-interested) or the common-interest (if she is common-interested).

***Success:*** Every voter acts according to the stated belief-desire pair.

That is, we can replace Self- Or Common-Interested Voting in the *extended argument from collective optimality* with the conjunction of the above three assumptions. These are, however, still rather strong assumptions. In this chapter, I mainly want to focus on one of these three, namely the Competence assumption that voters *correctly* judge the options by the lights of their desired ends — self- or common interest. Assuming the other two assumptions to be satisfied — that is, assuming every voter to be self- or common-interested and to vote according to their belief-desire pair, I want to explore how far Competence may be relaxed, while preserving the collective optimality of the weighted majority rule.

The arguments in this chapter build on a number of theorems that are generalisations and extensions of the well-known Condorcet jury theorem. I argue that these Condorcet theorems can be applied to the weighted majority rule under conditions of (empirical) uncertainty to show that this rule selects

the common-interest option with *near certainty*, even when Competence is not satisfied. The upshot is that the weighted majority rule can still be shown to be collectively optimal — though in a slightly different sense. Recall that we defined ‘being collectively optimal’ as ‘selecting the collectively optimal option’ (see e.g. premise (3) in the *original argument from collective optimality* in 3.2 above). We may now say that this definition gives us a *strong* notion of collective optimality. In this chapter, we are dealing with collective optimality in a weaker sense. Let us define ‘being weakly collectively optimal’ as ‘selecting the collectively optimal option with near certainty or certainty’. The argument in this chapter shows that the weighted majority rule is *weakly* collectively optimal. This conclusion is, obviously, weaker than the previous ones. It rests, however, on an improved argument that does not rely on the strong assumptions of (empirical) certainty and — in extension — of non-erratic voting behaviour.

I introduce the classical Condorcet jury theorem in 5.2.1. I then proceed to show how various generalisations and extensions of it are relevant to the present study. First, I consider only cases where everyone has equal stakes. I show that some of the Condorcet theorems apply to them, by arguing that they apply to cases where all voters are common-interested (5.2.2) and that this also holds when all voters are assumed to be self-interested (5.2.3). I then proceed to show that my arguments can be extended to ‘mixed’ cases with both self- and common-interested voters (5.2.4). These arguments presuppose that voter competences are distributed across the group in rather specific ways. In 5.2.5, I relax this assumption further by allowing any competence distribution. Second, I consider cases where stakes are unequal and where voters are thus assigned unequally sized bundles of votes. In 5.2.6, I show that a crucial assumption for the Condorcet theorems is violated in these cases, namely that of independence between votes. I then show that the dependence resulting from the voters indivisible vote bundles can be dealt with through yet another Condorcet theorem (5.2.7). In 5.2.8, I briefly consider an alternative approach to relaxing the independence assumption. Finally, in 5.3 I sum up and conclude.

Note that the Condorcet theorems I refer to have been developed and proved by other theorists. I utilise their theorems and proofs for my purposes. My main strategy is to show that the theorems' assumptions hold in the cases I discuss and hence that their results apply here as well. It is, however, not the task of the present study to analyse or question the theorems and proofs themselves. This means that my arguments are sound only to the extent that these theorems and proofs are correct.<sup>1</sup>

---

<sup>1</sup> I am indebted to Mats Ingelström for stimulating discussions on this chapter's subject.

## 5.2 The argument from weak collective optimality

Let us imagine a group of individuals who face a binary decision, whose stakes are defined in accordance with a given criterion of the common good (see 3.2 above). All the stake-holders ('voters') satisfy the Success assumption. Now, let us more specifically assume that all these stake-holders as it happens are common-interested, thereby complying with Self- Or Common-Interested Motivation. However, they do not correctly judge which option is in the common interest and thus fail to satisfy Competence. Then, Self- Or Common-Interested Voting does not follow, and thus we cannot use the *generic argument from collective optimality* in order to show that the weighted majority rule is collectively optimal for this case.

My task is now to build a new argument for this rule, which works for such cases with less than fully competent voters. That is, I want to state such an argument under the assumption that voters may be (empirically) uncertain as to which option promotes their self-interest or the common interest, respectively. There is an influential theorem regarding the results of *simple* majority rule with less than fully competent voters: the Condorcet jury theorem. I now state it and then apply it — in a series of steps — to the weighted majority rule.

To clarify some terminological issues: by 'voter competence' or 'individual competence' ( $c_i$ ) I refer to the probability that voter  $i$  correctly judges the options (according to a specified standard). If this probability is one ( $c_i = 1$ ), I call the voter 'fully competent'. If it is greater than chance ( $c_i > \frac{1}{2}$ ), I call her 'minimally competent'. If it is zero ( $c_i = 0$ ), I call her 'maximally incompetent'. By 'group competence' ( $P_n$ ) I refer to the probability that the majority of the votes is cast for the correct judgment.<sup>2</sup>

### 5.2.1 Introducing the Condorcet jury theorem

In its classical version, the Condorcet jury theorem applies to binary decisions between two propositions. Such a pair of propositions can be expressed, e.g. by the claim 'The defendant is guilty' and its negation. Likewise, it can be expressed by the claims ' $x$  (rather than  $y$ ) is in the common interest' and ' $y$  (rather than  $x$ ) is in the common interest'. It is assumed that only one of the claims in each pair is correct according to an independent

---

<sup>2</sup> I do not presuppose a specific interpretation of 'probability'. Note, though, that a subjectivist interpretation, identifying each voter's probability of a correct judgment with the degree of confidence that she herself assigns to her judgment, is not suitable here. However, a number of objectivist interpretations could be adopted e.g. a frequency interpretation (identifying the voter's probability of judging correctly with her relative frequency of correct judgments in a series of similar events) or a propensity interpretation (identifying this probability with the voter's tendency or disposition to judge correctly). What is important, though, is that the same interpretation is used consistently throughout the text.

standard (e.g. a standard of being guilty or of being in the common interest). Moreover, the following definitions are presupposed.

**Equal Minimal Competence:** Every voter is equally likely to judge the options correctly and somewhat more likely than not to judge the options correctly, such that every voter's individual competence is  $c > \frac{1}{2}$ .

**Voting According to Judgment:** Every voter votes according to this judgment.

**Voter Independence:** Every voter judges (probabilistically) independently, that is, how each judges the options does not depend on how others judge them.

The Condorcet jury theorem (CJT), in its classical form, then states:

**Classical CJT.** For binary decisions with exactly one correct option (according to some independent standard) and a group of  $n = 2m+1$  voters, given Equal Minimal Competence, Voting According to Judgment, and Voter independence, (i) the probability of a correct majority vote  $P_n$  is higher than any single voter's competence  $c$ , and (ii)  $P_n$  strictly increases with the number of voters  $n$  and approaches certainty as the number of voters increases to infinity.<sup>3</sup>

Some notes about this theorem: first, its scope is limited to odd-numbered groups of at least three voters. This limitation is common in the literature on the Condorcet jury theorem because it simplifies the exposition considerably. Still, it should be noted that the theorem can be extended to even-numbered cases of more than three voters.<sup>4</sup> Therefore, and in order to keep things simple, I only consider odd-numbered cases in the remainder of this chapter.

---

<sup>3</sup> Cf. e.g. Ladha (1992: 618), Miller (1986: 175) and Boland (1989: 182). Owen, Grofman and Feld (1989: 2) also make explicit the assumption that the 'prior odds as to which of the two alternatives is the correct one are even'. For a short proof of Classical CJT, see e.g. Ladha (1992: 632f.). In fact, Ladha's proof is conducted for any (even or odd) number  $n \geq 3$  of voters. Since some of the subsequent theorems are formulated for odd-numbered groups, I stick to this assumption from the outset for the sake of simplicity. For the original results, see Condorcet (1785).

<sup>4</sup> To be precise: for claim (ii), with an even-numbered group of  $n > 3$  voters, if ties are resolved by a random tie-breaker, the probability that a majority of these  $n$  voters votes for the correct option equals the probability that a majority of  $n-1$  voters votes for the correct option (cf. e.g. Miller 1986: 175; Ladha 1992: 618). That is, the probability of a correct majority vote is equivalent to that of the next-smaller odd-numbered group of voters. For claim (i), with an even-numbered group of  $n > 3$  voters, the probability of a correct majority vote is greater than any single voter's competence conditional on a stricter competence assumption. E.g. for  $n = 4$  voters, we need to assume  $c > 0.77$ ; for  $n = 12$ ,  $c > 0.56$ ; and for  $n = 102$ ,  $c > 0.5056$  (Bovens and Rabinowicz 2006: 4f.; 36). That is, voters must be significantly better than chance, but with a margin that decreases as the number of voters increases.

Second, the theorem presupposes some independent standard of correctness. This standard is independent in the sense of not causally depending on the individual votes and the judgments underlying them. To state a simple example: say that the independent standard is truth, and that a group of people votes for or against the proposition ‘There is life on Mars’. The truth value of this proposition is causally independent of the individuals' votes and underlying judgments concerning this issue. The same holds if the group votes on a proposition such as ‘The average height of all group members is below 1.70 meters’. The truth value of this proposition is causally dependent on the group members (to be precise: on their heights), but not on their votes and judgments concerning it.

Third, the theorem takes the form of a conditional, stating three conditions and a derived result. Concerning this result, claim (i) states that it is more likely that the majority votes for the correct option than that any single voter does. This is sometimes labelled the *non-asymptotic conclusion*. Claim (ii) states that this gets more and more likely (approaching certainty) as the group size increases (toward an infinite number of voters). This is sometimes labelled the *asymptotic conclusion*. I at times refer to these claims by these two labels. (Both claims are somewhat modified by the below extensions of the Condorcet jury theorem. For convenience, I keep the labels even for these slightly modified claims.)

Concerning the theorem's conditions, Equal Minimal Competence assumes that voters are equally (or homogeneously) better than chance in correctly judging the options by the lights of the independent standard. This is thus a considerably weaker assumption than the above employed Competence, which assumed correct judgments. (Note Equal Minimal Competence contains a further, implicit assumption, which I do not discuss in the following, and which holds for all considered Condorcet theorems below. The implicit assumption is that, for any binary decision with options  $x$  and  $y$ , voter competence  $c$  is the same, regardless of whether  $x$  or  $y$  in fact is the correct option.) In conjunction with Voting According to Judgment (which basically replaces the previous Success Assumption), Equal Minimal Competence implies that voters are better than chance in voting for the correct option. Voter Independence states that any voter  $i$ 's judgment is (probabilistically) independent of any other voter  $j$ 's judgment. This means that  $i$ 's probability of judging correctly is (probabilistically) independent of  $j$ 's probability of judging correctly. This can be spelled out more formally, in terms of  $i$ 's probability of judging correctly being equal to her conditional probability of judging correctly, given  $j$ 's probability of judging correctly. Alternatively, we can say that the probability that both  $i$  and  $j$  simultaneously judge correctly equals the product of their individual probabilities to judge correctly (that is, of their competences).

In order to get an intuitive understanding of the basic idea behind Classical CJT, consider David Estlund's illuminating illustration:

‘If you have 1.000 coins, with each one slightly weighted to turn up heads — say with a 51 percent chance — what is the chance that at least a majority of them will turn up heads? With that many coins, we know that very nearly 51 percent of them will turn up heads, and so it is quite likely indeed that more than 50 percent will’.<sup>5</sup>

The example, with 1.000 independent coin tosses, illustrates the above non-asymptotic conclusion (i). We can expand the example a bit: in comparison, if we tossed only three of these slightly weighted coins, the probability that at least a majority of them (two coins) would turn up heads would not be very high. But, if we had a million of these coins, it would seem extremely likely that at least a majority of them (500.001 coins) would land heads up. This illustrates the above asymptotic conclusion (ii).

The theoretical background of this phenomenon is the law of large numbers. It basically states that for large samples, the proportion of the numbers of items displaying a certain property will approximately correspond to the average probability for any single item to display the property in question — the larger the sample, the greater the degree of correspondence. Or, as List and Goodin aptly put it: ‘The point of the law of large numbers is that, although *absolute* deviations from the expected numbers still increase as the number of trials increases, those absolute deviations are a decreasing *proportion* of the total as the number of trials increases’.<sup>6</sup>

Note that the optimistic results of the classical Condorcet jury theorem have a sad flipside. If Equal Minimal Competence is exchanged for an *Equal Incompetence* assumption, such that every voter is assumed to be less likely than chance to judge correctly ( $c < \frac{1}{2}$ ), the results are reversed. This means that, for equally incompetent and independently judging voters who vote according to their judgment, (i\*) the probability of a correct majority vote  $P_n$  is lower than any single voter's competence, and (ii\*) this probability strictly *decreases* as the number of voters increases and approaches *zero* as the number of voters increases to infinity. Furthermore, if every voter is assumed to be exactly as competent as chance ( $c = \frac{1}{2}$ ), the probability of a correct majority vote  $P_n$  remains  $\frac{1}{2}$ , no matter how many voters there are.<sup>7</sup> I do not, however, subsequently explore these possibilities further, since I am interested in the conditions under which the weak collective optimality of the weighted majority can be preserved. Still, one should be aware of these possibilities throughout.

So, how is the Condorcet jury theorem supposed to apply to my present study? This is by no means obvious. Consider that Classical CJT presupposes, first, that there is one correct option, according to an *independent stand-*

---

<sup>5</sup> Estlund (2008: 15).

<sup>6</sup> List and Goodin (2001: 285).

<sup>7</sup> Cf. e.g. Ladha (1992: 618) and Miller (1986: 175). These ‘sad flipside’ results also apply to the below proposed generalisations and extensions of the Condorcet jury theorem.

*ard of correctness*. Second, it presupposes that all voters vote for what they judge to be the correct option, so implicitly, we take voters to be motivational truth- or *correctness-trackers*. Third, it presupposes that the decision is made by *simple* majority rule.

As a first objection against applying Classical CJT in the present study, one might then suggest that there is no correct option here, according to some independent standard of correctness. However, this is clearly not true.

Consider that we have assumed (see 2.3.2. above) that there is some criterion of the common good that determines which of the given options is in the common interest. It is not important *which* specific criterion we have accepted, only *that* we have accepted one. This criterion provides a standard according to which one of the options in a given binary decision can be evaluated. For a sum-total criterion, for instance, our independent standard is ‘being the option with at least as high sum-total of well-being as any other’. If we disregard cases where both options are equally good (since it does not matter, *ceteris paribus*, which of them is selected), this independent standard singles out, for any binary decision, exactly one option as the common-interest option. We may then say that this is the correct option, according to the independent standard of the accepted criterion.

This standard is of course not entirely independent of the voters, as e.g. the sum-total of well-being is constituted by the aggregate of the voters' well-being. But it is independent in the relevant sense since it does not causally depend on their votes and underlying judgments.<sup>8</sup> (Compare with the above example concerning a group's vote on their average height.)

As a second objection, one may claim that Classical CJT can only be applied to very specific cases: those where all voters judge the options in the light of the proposed independent standard — that is, those cases where all voters are common-interested. Yet we want to say something about cases with self-interested voters (or mixed cases) as well. The problem with self-interested voters is that they vote not according to their judgment of an independent standard of the common interest, which singles out one and the same option for all voters. Rather they vote according to their perception of their respective individual standard of self-interest, which may pick out one option as better for some of them, and the other option as better for others.

This is indeed a considerable complication. I deal with it in the following three sections. My argument, to the effect that Classical CJT is applicable, focuses first on groups of exclusively common-interested voters, is in a second step extended to hold for self-interested voters and in a third step for mixed groups of both kinds of voters.

---

<sup>8</sup> Note that the equal minimal competence assumption implies that these judgments are truth- or correctness-tracking, to some degree. Thus, there is *probabilistic* dependence between the independent standard and the voters' judgments. I return to different forms of dependence and their relevance for my arguments in 5.2.6 below.

As a third objection, one may note that Classical CJT is meant to apply to *simple* majority rule, while the present study deals with the *weighted* majority rule. Again, this is a complication I deal with below. Simple majority rule is a special instance of the weighted majority rule, for binary cases where everyone's stakes are equal. So I start out by only considering equal-stakes cases. For these cases, a number of Condorcet theorems can be directly applied to the weighted majority rule. I eventually extend the results to unequal-stakes cases, where simple and weighted majority rule differ (in their assignments of votes and hence, possibly, in their outcomes).

A fourth objection is that applying Classical CJT is not much of an improvement when it comes to the assumptions needed. Though this move allows us to relax the strong assumption of full voter competence, it brings with it pretty strong assumptions of its own, namely Equal Minimal Competence and Independence. In the course of this chapter, I explore how these new assumptions can be relaxed as well.

### 5.2.2 Equal-stakes cases with common-interested voters

Let us start then by considering cases where all voters have equal stakes and all are common-interested: their desired end is the promotion of the common interest. They are not motivational correctness-trackers whose desired end is to find out the correct option. This is, however, not an obstacle for applying Classical CJT. The precise nature of the voters' ends does not matter. What matters is that they, as common-interested voters, vote according to their judgment of the correct option, according to the independent standard of the common good (as assumed in Voting According to Judgment).

Hence, for these specific cases, Classical CJT can be employed as the first premise in a new argument for the weighted majority rule. As we will see shortly, the argument's conclusion (6) is simply an adaptation of Classical CJT to the present setting of equal-stakes cases with common-interested voters. This conclusion then specifies the conditions under which the weighted majority rule is weakly collectively optimal (that is, selects the common-interest option with near certainty or certainty) in these cases. Let us then call what follows the *first argument from weak collective optimality*. (Note that, just like the *generic argument from collective optimality*, it can be stated for any of the criteria of the common good I have discussed in 2.3.2 above.)

**(1) Classical CJT.** For binary decisions with exactly one correct option (according to some independent standard) and a group of  $n = 2m+1$  voters, given Equal Minimal Competence, Voting According to Judgment, and Voter independence, (i) the probability of a correct majority vote  $P_n$  is higher than any single voter's competence  $c$ , and (ii)  $P_n$  strictly increases with the number of voters  $n$  and approaches certainty as the number of voters increases to infinity.

(2) The correct option is the common-interest option, according to the independent standard of the given criterion of the common good.

(3) For equal-stakes binary decisions, the option for which a majority of voters votes is selected by the weighted majority rule (since the rule assigns equal votes to all voters and selects the option that receives a majority of votes).

(4) For any given level of equal minimal voter competence, there is a range of probability levels of a correct majority vote that we call ‘near certainty’ and (according to the Condorcet jury theorem) a corresponding range of numbers of voters  $n$  that we call ‘sufficiently large numbers’ of voters.

(5) A rule that with near certainty selects the common-interest option, as defined by the given criterion of the common good, is weakly collectively optimal, according to this criterion.

Hence:

(6) For equal-stakes binary decisions with exactly one common-interest option (according to the given criterion of the common good) and a group of a sufficiently large number  $n = 2m+1$  of common-interested voters, given Equal Minimal Competence, Voting According to Judgment, and Voter independence, the weighted majority rule is weakly collectively optimal (according to this criterion).

Thus, for equal-stakes cases with common-interested voters, the weighted majority rule can be shown to be weakly collectively optimal, without assuming that voters *infallibly* judge which option is in the common interest. Rather the weaker assumption that they are just equally better than chance at ‘guessing’ the common-interest option suffices — if only there are sufficiently large (odd) numbers of voters.

The argument works for any threshold of ‘near certainty’ that may be proposed and any given (minimal) competence level  $c > 1/2$ . For any such threshold and competence level  $c$ , one can compute the minimum number  $n$  of voters required to achieve the threshold. The lower the competence level, the more voters are needed to reach the threshold. Likewise, for any threshold of near certainty and a given number  $n$  of voters, one can compute a minimum level of competence  $c$  required to achieve the threshold. The fewer voters there are, the more competent they have to be to reach the threshold. (This latter point shows that not only brute numbers but also individual voter competence matters to the quality of democratic outcomes. Thus, the call for raising individual voter competence and thereby improving the democratic input — e.g. through education and deliberation — can be given an important place even in aggregative accounts of democratic decision-making; cf. 2.2.1 above.)

To precisify premise (4) of the above argument, and to get a grasp of how numbers and individual competence levels are correlated with group compe-

tence, consider *Table 1* (the entry in each cell gives the probability of a correct majority vote,  $P_n$ , for the stated values of the numbers of voters,  $n$ , and voter competence,  $c$ ).<sup>9</sup>

	$c = 0.51$	<b>0.55</b>	<b>0.6</b>	<b>0.75</b>	<b>0.9</b>	<b>0.95</b>
$n = 3$	0.5150	0.5748	0.6480	0.8438	0.9720	0.9928
<b>9</b>	0.5246	0.6214	0.7334	0.9510	0.9991	0.9999
<b>25</b>	0.5398	0.6924	0.8462	0.9981	0.9999	0.9999
<b>250</b>	0.6241	0.9440	0.9994	0.9999	0.9999	0.9999
<b>1.000</b>	0.7365	0.9993	0.9999	0.9999	0.9999	0.9999
<b>10.000</b>	0.9772	0.9999	0.9999	0.9999	0.9999	0.9999

*Table 1*

Note that the *first argument from weak collective optimality* does not employ all the information that is contained in its first premise, *Classical CJT*. The argument relies solely on the asymptotic-conclusion (ii). The non-asymptotic conclusion (i) is irrelevant throughout the argument. I choose to retain (i) in the above argument (as well as the others to come) because this conclusion is picked up again and used in another argument for the weighted majority rule (see 6.3 below).

### 5.2.3 Equal-stakes cases with self-interested voters

Now, what happens if we consider equal-stakes cases with *self-interested* voters? These cases differ in some important respects from the usual Condorcet cases with truth-tracking voters and from the above considered equal-stakes cases with common-interested voters. In the latter two scenarios, all voters are assumed to judge the options in the light of an independent standard (whether their desired end is to find the truth or to promote the common interest). Were they all fully competent (that is, certain) in the light of their desired end, and voted accordingly, all would then vote for the same option: the correct one. In these cases, the voters' competence level equals their probability to correctly judge (and vote for) the common-interest option.

In contrast, self-interested voters judge the option in accordance with their respective individual standard of self-interest. Were they all fully competent in the light of their desired end, and voted according to their judgment, some would vote for one option — and others for the other. In such a scenario, we know that the option that is in the self-interest of the majority stake-holders is in the common interest, while the option that is in the interest of the minority stake-holders is opposed to the common-interest option. Hence, given

---

<sup>9</sup> Cf. Miller (1986: 176), who also notes that ‘most ["0.9999"] entries actually round off to "1.0000" but this is not done to indicate residual uncertainty’.

Voting According to Judgment, fully competent majority stake-holders would in effect vote for the common-interest option, while the fully competent minority stake-holders would vote against it. And this means that the latter voters can be described as maximally *incompetent*, according to the independent standard of the common good. And likewise, a minority stakeholder who is slightly better than chance at correctly judging her self-interest option can be described as slightly worse than chance, according to the independent standard.

It is at this stage convenient to distinguish two notions of competence. On the one hand, we deal with the voters' competence in judging which option serves their own desired end — be it the promotion of self-interest or of the common interest. Let us call this a voter *i*'s *end-competence*,  $e_i$ . On the other hand, what matters to Classical CJT (and the other Condorcet theorems) is the voters' probability to judge correctly according to the independent standard — in our case: to judge correctly which option is in the common-interest. Let us call this a voter *i*'s *CJT-competence*,  $c_i$ . Now, a voter's CJT-competence may come apart from her end-competence, namely, whenever she is a self-interested minority member. Such a voter *j*'s CJT-competence equals her end-*incompetence*, that is,  $c_j = 1 - e_j$ . Thus, if she is maximally end-competent,  $e_j = 1$ , then she is maximally CJT-incompetent,  $c_j = 0$ .

Note that for any voter *i* who is not a self-interested minority stakeholder, her CJT-competence equals her end-competence, that is,  $c_i = e_i$ . If she is common-interested, her desired end is the promotion of the common interest, and hence her end-competence equals her CJT-competence. And if she is a majority stakeholder, her self-interest coincides with the common interest, and hence, whatever her desired end among these two, her end-competence equals her CJT-competence.

The connection between the two notions of competence can be illustrated with this simple example. Let us say that all voters in a group are self-interested and fairly competent when it comes to correctly judging which option is in their self-interest, such that for each voter *i*, her end-competence  $e_i = 0.75$ . Then, each majority member *i* has a CJT-competence of  $c_i = 0.75$ , as her self-interest coincides with the common interest. Yet each minority member *j* only has a CJT-competence of  $c_j = 0.25$ , since the option that is in her self-interest is opposed to the common-interest option. We could thus describe the minority stake-holders as worse than chance at judging what is in the common interest — although, and because, they are better than chance at judging what is in their self-interest.

(Note that in the following text, whenever necessary, I make explicit whether CJT-competence or end-competence are at play. Whenever no such specification is made, e.g. when referring to the Condorcet literature, 'competence' refers to 'CJT-competence'.)

All this means that, if we, as before, assume equal minimal competence and spell this out as equal minimal *end-competence*, the voters' CJT-

competence may vary and may be way worse than chance, as in the just considered illustration.<sup>10</sup> That is, for self-interested voters, assuming equal minimal end-competence implies that voters may be heterogeneously and less than minimally (CJT-)competent. This, however, contradicts Classical CJT's assumption of Equal Minimal Competence, which means that the theorem cannot be applied in the present cases.

However, the Condorcet jury theorem has been generalised to cover cases with heterogeneous and less than minimal voter competence. My following argument establishes that such a generalisation does apply to the present cases.

The argument proceeds in two steps. First, I establish a partial *Heterogeneous competence CJT*, which gives us the asymptotic conclusion (ii) that the probability of a correct majority vote strictly increases with the number of voters and approaches certainty as the number of voters increases to infinity. Second, I establish another partial *Heterogeneous competence CJT*, which gives us a non-asymptotic conclusion (which is slightly different from the above claim (i) that the probability of a correct majority vote is greater than any single voter's competence). The reason for this argumentative structure is that these two theorems have been stated and proved separately.

Owen, Grofman and Feld show that the asymptotic conclusion (ii) holds for heterogeneously (CJT-) competent voters whenever their *average* (CJT-) competence is greater than chance.<sup>11</sup> We can define the following.

***Minimal Average CJT-Competence:*** The voters are minimally (CJT-)competent on average, that is, the average of their individual probabilities to judge the options correctly (according to the independent standard) is above chance,  $c^* > \frac{1}{2}$ .

Then, we can state the theorem as follows.

***Heterogeneous Competence CJT (Part 1):*** For binary decisions with exactly one correct option (according to some independent standard) and a group of  $n = 2m+1$  voters, given Minimal Average CJT-Competence, Voting According to Judgment, and Voter independence, (ii) the probability of a correct majority vote  $P_n$  strictly increases with the number of voters  $n$  and approaches certainty as the number of voters increases to infinity.

---

<sup>10</sup> I say that CJT-competence *may* vary because there are cases where it does not: namely, when there are no minority members, that is, when one and the same option is in every voter's self-interest. Note that we could simply assume all voters to be equally minimally CJT-competent, and then proceed to apply Classical CJT and make the argument from weak collective optimality. This, however, would be an overly artificial assumption since it would imply that all majority stake-holders are equally better than chance and all minority stake-holders equally worse than chance (by the same margin) at judging their self-interest option. It is also an unnecessary assumption, given that we can solve the problem in another way, as shown in the remainder of this section.

<sup>11</sup> Owen, Grofman and Feld (1989: 3f.).

The following argument shows that this theorem can be applied to the present case of self-interested voters (with equal stakes), if we assume that they are equally minimally end-competent. First, I show that Minimal Average CJT-Competence is implied by the following assumption.

**Equal Minimal End-Competence:** Every voter is equally better than chance at judging the options correctly in the light of her desired end, such that every voter's individual end-competence is  $e > \frac{1}{2}$ .

Note that *average CJT-competence*,  $c^*$ , is simply the sum of all voters' CJT-competences divided by the number of voters. Let us look at a group of  $n$  voters who satisfy Equal Minimal End-Competence. Their end-competence is better than chance, which can be expressed as  $e = \frac{1}{2} + b$  (with  $0 < b \leq \frac{1}{2}$ ). Within the group, there is a majority of  $m = n/2 + a$  voters  $i$  (with  $0 < a \leq n/2$ ) whose CJT-competence equals their end-competence (thus,  $c_i = \frac{1}{2} + b$ ). Moreover, there is a minority of  $n - m = n/2 - a$  voters  $j$  whose CJT-competence equals their end-*in*competence (this means that  $c_j = \frac{1}{2} - b$ ). Hence, the average CJT-competence  $c^*$  is:

$$\begin{aligned} c^* &= (mc_i + (n-m)c_j)/n \\ &= [(n/2+a)(\frac{1}{2}+b) + (n/2-a)(\frac{1}{2}-b)]/n \\ &= (n/4 + a/2 + nb/2 + ab + n/4 - a/2 - nb/2 + ab)/n \\ &= (n/2 + 2ab)/n \\ &= \frac{1}{2} + 2ab/n. \end{aligned}$$

Since  $a$ ,  $b$ , and  $n$  all are greater than 0, average CJT-competence  $c^*$  must be strictly greater than  $\frac{1}{2}$ , that is, above chance. This proves that for all here considered equal-stakes cases with self-interested voters, Equal Minimal End-Competence implies Minimal Average CJT-Competence. Hence, given Voting According to Judgment and Voter Independence, Heterogeneous Competence CJT (Part 1) can be applied in the here considered equal-stakes cases with self-interested, equally minimally competent voters.<sup>12</sup>

This theorem does, however, only state the asymptotic conclusion (ii). It does not give us a non-asymptotic conclusion. Now, note that the latter, as stated in Classical CJT above, claims that the probability of a correct majority vote exceeds *any single* voter's competence. In cases with heterogeneously

---

<sup>12</sup> Note that Voting According to Judgment implies, for a self-interested minority stake-holder  $j$  with end-competence  $e_j = 0.75$ , and hence CJT-competence  $c_j = 0.25$ , that this voter has a 0.75 chance to vote for her self-interest option (as she intended) and a 0.25 risk of voting for the common-interest option instead. And likewise, for a self-interested majority stake-holder  $i$  with end-competence  $e_i = 0.75$ , and hence CJT-competence  $c_i = 0.75$ , this assumption means that she has a 0.75 chance to vote for her self-interest option (which coincides with the common-interest option), and a 0.25 risk of voting against it by mistake. Thus this assumption does not spell trouble for cases with self-interested voters who vote according to different ends.

competent voters (who are assumed to be minimally CJT-competent on average), some single voter  $i$ 's competence  $c_i$  could be as high as 1. Then, the non-asymptotic conclusion would be an impossible order. So it makes more sense to look for a theorem that gives us a non-asymptotic conclusion reformulated in terms of the voters' *average* competence, such as:

(i') the probability of a correct majority vote  $P_n$  is higher than the voters' average (CJT-)competence  $c^*$ .

Now, (i') is easily shown to be false for certain cases, even when average competence is above chance. For instance, consider a group of three voters, one of whom has a competence of  $c_1 = 1$ , while the other two have a competence of  $c_2 = c_3 = 0.28$ . Their average competence is  $c^* = 0.52$ . Yet the probability that a majority votes for the common-interest option is only 0.4816, that is, below average competence.<sup>13</sup> Thus, (i') does not hold.

So what, then, is going on in such cases? It is, simply, that competence is distributed in such a way that a highly competent minority can push the average competence of the group above chance, while a worse-than-chance majority pushes the probability of a correct majority vote below average competence. But let us also consider the opposite case, with a highly competent majority and a worse-than-chance minority. Here, average competence may even be below chance — while the probability of a correct majority vote is above chance. Consider a group of three where one voter has a competence of  $c_1 = 0$ , while the other two have a competence of  $c_2 = c_3 = 0.72$ . Then, their average competence is  $c^* = 0.48$ . Yet the probability of a correct majority vote is 0.5184 — above chance, and above average competence.<sup>14</sup>

The impact of the distribution of individual competences on the probability of a correct majority vote is captured in yet another Condorcet theorem. Ladha shows that, if individual competences are distributed such that the

---

<sup>13</sup> This figure can be easily calculated as the sum of the probabilities that — while the first voter certainly votes for the common-interest option — the second votes for and the third against, or the second against and the third for, or both the second and the third vote for the common-interest option. That is:  $1 \cdot 0.28 \cdot 0.72 + 1 \cdot 0.72 \cdot 0.28 + 1 \cdot 0.28 \cdot 0.28 = 0.4816$ . This 'puzzling' example is taken from Grofman, Owen and Feld (1983: 271), who draw our attention to the 'quite counterintuitive result that [...] a group can have [average voter competence  $p^*$ ]  $> \frac{1}{2}$  and yet have [group competence]  $PN < \frac{1}{2}$ . For example: (a) (1, 0.28, 0.28); [ $p^*$ ] = 0.52, yet  $PN = 0.4816$ . (b) (1, 1.0, 0.2, 0.2, 0.2); [ $p^*$ ] = 0.52 yet  $PN = 0.488$ '. (However, the larger the group, the less likely that such puzzling cases occur — for a formal explanation of this latter fact, see Boland's (1989) results stated below in 5.24.)

<sup>14</sup> This figure can be easily calculated as the probability that — while the first voter certainly votes against the common-interest option — the second and third voters vote for the common-interest option. That is:  $0.72 \cdot 0.72 = 0.5184$ . This example is again from Grofman, Owen and Feld (1983: 271), who state that 'a group can have [ $p^*$ ]  $< \frac{1}{2}$  and yet have  $PN > \frac{1}{2}$ . For example: (a) (0.72, 0.72, 0); [ $p^*$ ] = 0.48, yet  $PN = 0.5184$ . (b) (0.8, 0.8, 0.8, 0, 0); [ $p^*$ ] = 0.48,  $PN = 0.512$ . (c) (0.8, 9.0, 0.7, 0, 0); [ $p^*$ ] = 0.48,  $PN = 0.504$ '.

probability that at least a majority ( $m = (n+1)/2$ ) of voters judges correctly is higher than the probability that at most a minority ( $n-m$ ) of voters judges correctly (given, we may add, Voting According to Judgment and Voter Independence), then the probability of a correct majority vote is greater than the voters' *average* (CJT-)competence.<sup>15</sup> We can thus define:

**Competence Distribution:** The probability that at least  $m = (n+1)/2$  voters judge correctly (according to the independent standard) is higher than the probability that at most  $n-m$  voters judge correctly.

Ladha's theorem can then be stated as follows.

**Heterogeneous competence CJT (Part 2):** For binary decisions with exactly one correct option (according to some independent standard) and a group of  $n = 2m+1$  voters, given Competence Distribution, Voting According to Judgment, and Voter Independence, (i') the probability of a correct majority vote  $P_n$  is higher than the voters' average (CJT-)competence  $c^*$ .

Now, can this theorem be applied to the here considered equal-stakes cases with self-interested voters? In these cases, there is a majority of at least  $m = (n+1)/2$  voters and a minority of at most  $n-m$  voters. As stated above, given Equal Minimal End-Competence, everyone's end-competence is better than chance, which can be expressed as  $e = 1/2+b$  (with  $0 < b \leq 1/2$ ). Since every majority member  $i$ 's CJT-competence equals her end-competence,  $c_i = 1/2+b$ . And since every minority member  $j$ 's CJT-competence equals her end-incompetence,  $c_j = 1/2-b$ . Hence, the majority members are more competent, that is, more probable to judge correctly than the minority members. And this means that it is more probable that the majority (of at least  $m = (n+1)/2$  voters) judges correctly than that the minority (of at most  $n-m$  voters) does. In other words, for the here considered equal-stakes cases with self-interested voters, if they satisfy Equal Minimal End-Competence, they satisfy Competence Distribution. Thus (given also Voting According to Judgment and Voter Independence), Heterogeneous competence CJT (Part 2) can be applied to the present equal-stakes cases with self-interested voters.

Now, for convenience, we can combine the two partial Heterogeneous competence CJTs, such that:

---

<sup>15</sup> Ladha (1993: 77). Cf. even Berend and Sapir (2007: 515) who spell out this condition in terms of 'that there is a bias towards having more members voting correctly rather than incorrectly'. It is easily checked that Ladha's condition is *not* satisfied in the first of the just considered three-voter cases (with  $c_1 = 1, c_2 = c_3 = 0.28$ ; cf. footnote 8), while it is satisfied in the second (with  $c_1 = 0, c_2 = c_3 = 0.72$ ; cf. footnote 9). For further work on the implications of the distribution of individual competence on group competence, see e.g. Grofman, Owen and Feld (1983: 268), Grofman (1978: 51–52) and Berend and Paroush (1998: 483).

***Heterogeneous Competence CJT:*** For binary decisions with exactly one correct option (according to some independent standard) and a group of  $n = 2m+1$  voters, given Minimal Average CJT-Competence, Competence Distribution, Voting According to Judgment, and Voter independence, (i') the probability of a correct majority vote  $P_n$  is higher than the voters' average (CJT-) competence  $c^*$ , and (ii) the probability of a correct majority vote  $P_n$  strictly increases with the number of voters  $n$  and approaches certainty as the number of voters increases to infinity.

This theorem can now be employed as a first premise in the *second argument from weak collective optimality*. The second premise is a summary of my just stated arguments.

(1) *Heterogeneous Competence CJT.*

(2) For equal-stakes cases with self-interested voters, given Equal Minimal End-Competence, then Minimal Average CJT-Competence and Competence Distribution.

(3) The correct option is the common-interest option, according to the independent standard of the given criterion of the common good.

(4) For equal-stakes binary decisions, the weighted majority rule selects the option for which a majority votes (since it assigns equal votes to all voters and selects the option which receives a majority of votes).

(5) For any given level of average (CJT-)competence (derived from a set of individual end-competences), there is a range of probability levels of a correct majority vote that we call 'near certainty' and a corresponding range of numbers of voters that we call 'sufficiently large' numbers.

(6) A rule that with near certainty selects the common-interest option, as defined by the given criterion of the common good, is weakly collectively optimal, according to this criterion.

Hence:

(7) For equal-stakes binary decisions with exactly one common-interest option (according to the given criterion of the common good) and a group of a sufficiently large number  $n = 2m+1$  of self-interested voters, given Equal Minimal End-Competence, Voting According to Judgment, and Voter Independence, the weighted majority rule is weakly collectively optimal (according to this criterion).

Thus, for equal-stakes cases with self-interested voters, the weighted majority rule can be shown to be weakly collectively optimal, without assuming that voters *correctly* judge what is in their self-interest. Rather, the weaker assumption that they are just equally better than chance at 'guessing' their

self-interest option suffices — if there is a sufficient (odd) number of voters.<sup>16</sup>

#### 5.2.4 Equal-stakes mixed-motivation cases

Thus far, we have considered equal-stakes cases with, we might say, homogeneously motivated voters: all are either assumed to be exclusively common-interested or exclusively self-interested. What happens if there is a mix of self- and common-interested voters in the group? Consider the following series of scenarios with heterogeneously motivated voters.

**Only self-interested voters.** In this equal-stakes binary decision, each voter  $i$  is self-interested and equally minimally end-competent, with  $e_i > \frac{1}{2}$ . From this latter assumption, we can infer Minimal Average CJT-Competence and Competence Distribution (as shown in the previous section). Then, as shown by the *second argument from weak collective optimality* above, given Voting According to Judgment and Independence, the weighted majority rule is weakly collectively optimal for sufficiently large (odd) numbers of voters.

**One common-interested voter.** Now, we change the above scenario just a bit, such that all voters — except one — are self-interested. This voter,  $j$ , is instead common-interested and (as before) minimally end-competent with  $e_j > \frac{1}{2}$ . That is, her desired end has changed from self-interest to the common interest, while her end-competence is the same as before. Then, there are two possibilities.

Either (1)  $i_j$  is a majority stake-holder. This means that the option that is in her self-interest is also in the common interest. Then, the present scenario is extensionally equivalent to the previous one, such that we can infer the weighted majority rule's weak collective optimality, given Voting According to Judgment, Independence, and sufficiently large (odd) numbers of voters.

---

<sup>16</sup> The argument in this section has been devised independently of a similar argument by Miller (1986). Miller considers binary decisions for voters with 'conflicting interests' and defines as an independent standard of success for the collective decision that 'the interests of the majority prevail' (1986: 178). He then shows that, for an odd number of  $n \geq 3$  voters, if all voters perceive their individual interest with probability  $0.5 < p < 1$ , then '(i)  $P'_n$  [the probability that majority interests prevail]  $> p^*$  [the expected proportion of the vote in favor of the majority position]; (ii)  $P'_n$  increases as  $n$  increases; and (iii)  $P'_n \rightarrow 1$  as  $n \rightarrow \infty$ '. (1986: 180). Since Miller's expression ' $0.5 < p < 1$ ' is equivalent to my 'equal minimal end-competence', his ' $P'_n$ ' is equivalent to my 'probability of a correct majority vote', and his ' $p^*$ ' is equivalent to my 'voters' CJT-competence on average', his extension of the Condorcet jury theorem is in fact equivalent to the conjunction of the above Heterogeneous competence CJT and Equal Minimal End-Competence.

I have refrained from replacing my two-step argument, involving Heterogeneous competence CJT Part 1 and 2, with Miller's more straightforward theorem. The reason is that I hope that my elaborations around the Equal Minimal End-Competence assumption, in terms of its implications on the distribution and group average of CJT-competence (the Minimal Average CJT-Competence and Competence Distribution assumptions), which result in the argument's second premise, can make the results more intuitively intelligible.

Or (2)  $i_j$  is a minority stake-holder. This means that the option that is in her self-interest is not in the common interest. Then, the present scenario differs from the one above in the following respect. In the *Only self-interested voters* case, since minority stake-holder  $j$  is self-interested, her CJT-competence equals her end-*in*competence. That is,  $c_j = 1 - e_j < \frac{1}{2}$ . Yet in the present *One common-interested voter* case, since she now is common-interested,  $j$ 's CJT-competence equals her end-competence. That is,  $c_j = e_j > \frac{1}{2}$ , which means that, compared to the previous scenario,  $j$ 's CJT-competence is now higher. Thus, the group's average CJT-competence must be higher. Thus, in this case, Minimal Average CJT-Competence must still hold.

Concerning Competence Distribution, note that we currently assume  $j$  to be a minority stake-holder. Yet, as a common-interested voter, she votes as if she were a majority stake-holder. Hence, we can treat her *as if* she were an additional majority stake-holder. Then, if Competence Distribution holds in the previous scenario with  $m > \frac{1}{2}$  majority stake-holders, it must hold as well for the present case, where there are, as it were,  $m+1$  majority stake-holders. Thus, in this case, Competence Distribution must still hold.

So even in this case, given Voting According to Judgment and Independence, Heterogeneous competence CJT can be applied. Adding the other premises (3) to (6) from the above *second argument from weak collective optimality*, we can then again infer that the weighted majority rule is weakly collectively optimal, given sufficiently large (odd) numbers of voters.

We can now proceed to *Two common-interested voters, Three..., Four...* and so on, all the way to *Only common-interested voters*, where all voters are common-interested. Every time, the same kind of reasoning applies. For each case, given Voting According to Judgment and Independence, Heterogeneous Competence CJT can be applied and, adding premises (3) to (6) from the above *second argument from weak collective optimality*, the weighted majority rule can be shown to be weakly collectively optimal for sufficiently large (odd) numbers of voters. This means that the above *second argument* can be adapted to a *third argument from weak collective optimality*, for 'mixed' groups of self- and common-interested voters, such that ultimately the following conclusion is established.

**Conclusion:** For equal-stakes binary decisions with exactly one common-interest option (according to the given criterion of the common good) and a group of a sufficiently large number  $n = 2m+1$  of common- or self-interested voters, given Equal Minimal End-Competence, Voting According to Judgment, and Voter Independence, the weighted majority rule is weakly collectively optimal (according to this criterion).

Thus, for equal-stakes 'mixed' cases with self-interested or common-interested voters, the weak collective optimality of the weighted majority rule can be established without assuming that voters *correctly* judge what is in their self-interest or in the common interest. Rather, the weaker assumption that they are just equally better than chance at 'guessing' their respec-

tive self- or common-interest option suffices — if there are sufficiently large (odd) numbers of voters.

### 5.2.5 ‘Mixed’ cases with heterogeneously end-competent voters

It may now be objected that Equal Minimal End-Competence is still a pretty strong assumption. Could we rebuild the argument from weak collective optimality on the much weaker assumption of *heterogeneous* (and possibly even *worse-than-chance*) end-competence?

Relaxing this assumption comes at the price that we no longer can infer Minimal Average CJT-Competence and Competence Distribution. Hence, we can no longer apply Heterogeneous Competence CJT, and thus no longer make the *second* (or adapted *third*) *argument from weak collective optimality*.

Of course, we could simply assume that Minimal Average CJT-Competence and Competence Distribution hold — not because Equal Minimal End-Competence holds, but for some other reason. However, as Competence Distribution is a rather strong assumption in its own right, it is tempting to try to make do without it. I show in the following sections that we can restate the argument from weak collective optimality from another Condorcet theorem, which does not rely on any specific distribution of individual competences.

Philip J. Boland (1989: 183) shows that the non-asymptotic and asymptotic conclusions (i') and (ii) hold for voters with any distribution of competence levels, whenever their *average* competence exceeds chance by at least  $1/2n$ .<sup>17</sup> We can thus define:

**Raised Average CJT-Competence:** The voters' average (CJT-)competence  $c^*$  exceeds chance by at least  $1/2n$ , that is,  $c^* \geq \frac{1}{2} + 1/2n$ .

Boland's theorem can now be restated as follows:

**Distribution-Neutral CJT.** For binary decisions with exactly one correct option (according to some independent standard) and a group of  $n = 2m+1$  voters, given Raised Average CJT-Competence, Voting According to Judgment,

---

<sup>17</sup> Boland (1989) makes explicit that (ii) applies for a *fixed* average competence  $c^*$ . This simply means that the strict increase in the probability for a correct majority vote is shown to hold for increasing group size when average competence is held constant. (It might not hold for a larger group with a lower average competence, even if this average competence is still above the required level. Or to put it differently: the theorem holds for groups of increasing size if adding voters does not diminish average competence.) This condition is arguably implicit in the previous theorems Heterogeneous competence CJT Part 1 and Part 2.

On a different note: Boland's (1989) competence condition  $c^* \geq \frac{1}{2} + 1/2n$  is not satisfied in the above ‘puzzling’ examples (cf. footnote 12 above) and thus helps explain them. (It does not, however, explain the examples referred in footnote 13 above.)

and Voter Independence, (i') the probability of a correct majority vote  $P_n$  is higher than the voters' average (CJT-)competence  $c^*$ , and (ii) the probability of a correct majority vote  $P_n$  strictly increases with the number of voters  $n$  and approaches certainty as the number of voters increases to infinity.

As we can see here, getting rid of Competence Distribution — in fact, getting rid of Equal Minimal End-Competence by which it was implied — comes at the price of a higher threshold for average CJT-competence. How high this price is depends on the value of  $n$ , that is, on the group size. The value of  $\frac{1}{2} + \frac{1}{2n}$  approaches  $\frac{1}{2}$  as the number of voters approaches infinity. This means that for sufficiently large groups, average common-interest competence just needs to be minimal, that is, only slightly above chance.

Let us have a look at what this somewhat stricter competence assumption means for our present equal-stakes cases. The average CJT-competence is simply the sum of all voters' CJT-competence levels, divided by the number of voters. A voter's CJT-competence equals her end-competence if she is a majority stake-holder or a common-interested minority stake-holder, and equals her end-*in*competence if she is a self-interested minority stake-holder. This can be seen when considering that a self-interested minority stake-holder's desired end is opposed to the common interest, while all the others desired end is in the common-interest (recall that for majority stake-holders self- and common-interest coincide). Thus, from the voters' end-competence levels, we can calculate the group's average CJT-competence level, which, as now stated, must be at least  $\frac{1}{2} + \frac{1}{2n}$ .

We can then formulate a *fourth argument from weak collective optimality*, with Distribution-Neutral CJT as its first premise, as follows.

(1) *Distribution-Neutral CJT.*

(2) For equal-stakes cases with self- or common-interested voters, if the voters have any end-competences such that their average (CJT-)competence is at least  $\frac{1}{2} + \frac{1}{2n}$ , then Raised Average CJT-Competence holds.

Then, adding premises (3) to (6) from the above *second argument from weak collective optimality*, we can establish the following conclusion.

**Conclusion:** For equal-stakes binary decisions with exactly one common-interest option (according to the given criterion of the common good) and a group of a sufficiently large number  $n = 2m+1$  of common- or self-interested voters, given end-competences such that Raised Average CJT-Competence holds, and given Voting According to Judgment, and Voter Independence, the weighted majority rule is weakly collectively optimal (according to this criterion).

This means that the weighted majority can be shown to be collectively optimal for all equal-stakes cases with self- *or* common-interested voters, with

different (and even less-than-minimal) end-competences, for sufficiently large (odd) numbers of voters, and given Raised Average CJT-Competence, Voting According to Judgment, and Voter Independence.

Note that, in fact, formulating introducing Raised Average CJT-Competence in effect removes the need for any specific assumption concerning voter motivation: voters may be motivated to vote according to any common-, self-, or partial-interested desired end. As long as their individual CJT-competence — as calculated from their individual end-competence in the light of their specific end — does not bring down average CJT-competence below the given limit of  $\frac{1}{2} + 1/2n$ , the *fourth argument from weak collective optimality* applies.

Also note that since *simple* majority rule is extensionally equivalent to the weighted majority rule in the here considered equal-stakes cases, these results apply to it as well. This observation brings us back to Wolff's alleged 'mixed motivation' problem (as described in 4.2.1 above). Recall that Wolff presupposes simple majority rule and purports to show that 'if part of the electorate vote in pursuit of their own interests, and part for common good [defined as majority interest], then it is possible to arrive at a majority decision which is neither in the majority interest, nor believed by the majority to be for the common good'.<sup>18</sup> And recall that I observed above that this formulation of the conclusion obfuscated the real problem: not that there is mixed motivation, but that there is epistemic failure.

Now that we have an analytical handle on voter incompetence (or epistemic failure), let us briefly reconsider Wolff's example.

**Mixed motivation.** There is a group facing two options,  $x$  and  $y$ .  $x$  is in the self-interest of 40% of the group (I call this the first subgroup), while  $y$  is in the self-interest of 60% (the second subgroup). Moreover, 80% of the entire group (across the first and second subgroup) believe  $y$  to be the common good, defined as majority interest, while 20% believe this of  $x$ . Now, suppose that all voters within the first subgroup vote according to their self-interest, while all the voters within the second vote according to their perception of the common interest. Then, 52% will vote for  $x$  — the option that is *not* in the majority interest and *not* believed to be so by a majority of the group. (This can be easily seen, as all in the first subgroup — 40% of the group — will vote for  $x$ , as well as the 20% of those within the second subgroup — constituting 60% of the group — who take  $x$  to be the common good, which amounts to an additional 12%.)

What the example assumes, in the terminology of this chapter, is that there are 40% minority stake-holders  $i$  who are self-interested and maximally end-competent, with  $e_i = 1$ . This gives them a CJT-competence of  $c_i = 0$ . Moreover, there are 60% majority stake-holders who are common-interested. Among these, an 80% subset of voters  $j$  are maximally end-competent, with

---

<sup>18</sup> Wolff (1994: 194).

$e_j = 1$ , such that their CJT-competence is  $c_j = 1$ , and a 20% subset of voters  $k$  are maximally end-incompetent, with  $e_k = 0$ , such that their CJT-competence is  $c_k = 0$ . This means that their average CJT-competence  $c^* = 40\% \cdot 0 + 60\% \cdot 80\% \cdot 1 + 60\% \cdot 20\% \cdot 0 = 0.48$ . As this is well below  $\frac{1}{2} + 1/2n$  (in fact, even below chance), the poor result should not come as a big surprise.

Let us return to our present case for the weighted majority rule under less than full voter competence. Until now we have only considered equal-stakes cases, in order to simplify the investigation. The next question is whether we can broaden the attained results to cover unequal-stakes cases as well.

### 5.2.6 Unequal-stakes cases and the independence assumption

There is one seemingly promising strategy to straightforwardly apply all of the above results to unequal-stakes cases, in which voters may have different numbers of votes. The suggestion is to simply replace the term ‘voter’ with the term ‘vote’ in the above arguments, or, to put it differently, to treat every vote the weighted majority rule assigns in unequal-stakes cases as if it were a separate voter.

To illustrate this suggestion: assume that there are three voters,  $i_1$ ,  $i_2$  and  $i_3$ , and that  $i_1$ 's stakes are three times as high as  $i_2$ 's and  $i_3$ 's, respectively. The weighted majority rule then assigns, say, three votes to  $i_1$  and one vote each to  $i_2$  and  $i_3$ . Thus, there are five votes distributed among the three voters. We could now treat this case just as a case with five voters, where each voter has one vote, and proceed with the argument as before.

In other words, the suggested strategy is that we restate the above conclusion of the *fourth argument from weak collective optimality* in terms of *votes* rather than *voters*, since what matters for the weighted majority rule in unequal-stakes cases is the number (majority) of votes, rather than of voters. That is, we do not focus on common- or self-interested voters  $i$ , who cast their votes for the intended option with any end-competences  $e_i$ , such that their average CJT-competence  $c^* \geq \frac{1}{2} + 1/2n$ . Instead we focus on votes  $v$ , that are cast according to the voters' judgment with any end-competences  $e_i$ , such that the votes have an *average probability*  $p^* \geq \frac{1}{2} + 1/2n$  of being cast for the common-interest option. We can thus define:

***Raised Average CJT-Probability:*** All votes are cast according to the voters' judgments, with any end-competences, such that the votes' average probability of being cast for the common-interest option  $p^*$  exceeds chance by at least  $1/2n$ , that is,  $p^* \geq \frac{1}{2} + 1/2n$ .

We then also have to redefine Voter Independence. In analogy to my characterisation of *voter* independence (see 5.2.1 above), independence between votes can be spelled out in terms of that *vote*  $v_i$ 's probability of being cast for the correct option equals its conditional probability of being so cast, given

vote  $v_2$ 's probability of being cast for the correct option. Alternatively, the probability that both  $v_1$  and  $v_2$  are simultaneously cast for the correct option equals the product of their individual probabilities. Let us then define:

**Vote Independence:** Every vote is cast probabilistically independently, that is, how each vote is cast does not depend on how others are cast.

Then we could suggest the following.

**Suggested Conclusion:** For equal-stakes binary decisions with exactly one common-interest option (according to the given criterion of the common good) and a group of a sufficiently large number  $n = 2m+1$  of votes that are cast by any number  $l \leq n$  of voters, given Raised Average CJT-Probability and Vote Independence, the weighted majority rule is weakly collectively optimal (according to this criterion).<sup>19</sup>

However, we can now see that the Suggested Conclusion runs into a big problem. Recall that the weighted majority rule assigns to each voter a number of votes, as an *indivisible* vote bundle. But then, if some voters have indivisible vote bundles with more than one vote — as would be the case with unequal stakes — Vote Independence does not hold.

To illustrate the problem: say that we know that common-interested voter  $i_1$  is given an indivisible bundle of two votes, and that she is minimally end-competent. We can infer that each of her two votes is slightly more likely than not to be cast in the common interest. But if we learn that one of  $i_1$ 's votes has been in fact cast in the common interest, then we can infer a change in probability, namely that  $i_1$ 's other vote has been cast in the common interest *with certainty*. In other words, the probability for one of  $i_1$ 's votes  $v_1$  to be cast in the common interest does *not* equal  $v_1$ 's conditional probability of being cast in the common interest, given that  $v_2$  is cast in the common interest. Thus,  $i_1$ 's votes  $v_1$  and  $v_2$  are not independent.

How damaging is such a violation of the independence condition? Before we can answer this question, we need to be clear about which kind of independence is violated in unequal-stakes cases, and which kind is assumed in the stated Condorcet theorems.

In the literature on the Condorcet jury theorem, violations of the independence assumption are usually assessed in terms of *positive* probabilistic correlation of judgments or votes: that is, that a pair of judgments or votes is *more* likely to concur than probabilistically expected.<sup>20</sup> Such correlation may

---

<sup>19</sup> Note that Voting According to Judgment is implied by Raised Average CJT-Probability, as the latter states that votes are cast according to the voters' judgment.

<sup>20</sup> Correlation between judgments or votes may also be *negative*, namely, when a pair of judgments or votes is *less* likely to concur than probabilistically expected. I start out by focusing on positive correlation, as arising from indivisible vote bundles. I do, however, return to negative correlation in the next section.

be due to causal connections between the votes themselves — let us call this *direct causal dependence* — or between the votes and other factors that function as common causes — let us call this *common cause dependence*.<sup>21</sup> Direct causal dependence between votes may, e.g. result when the vote of a ‘voting leader’ influences the votes of other voters.<sup>22</sup> Common cause dependence may, e.g. result when voters take part in mutual deliberation, or belong to the same ‘schools of thought’, when such partaking or affiliation brings them to perceive things in similar ways.<sup>23</sup> Another common cause of probabilistic correlations between votes is commonly shared evidence.<sup>24</sup> Even seemingly irrelevant factors, such as the weather, might work as a common cause. Sunny skies, for instance, might bring voters to look more optimistically at a pair of options and to misjudge their harmful effects.<sup>25</sup>

The independence assumption of all the above stated Condorcet theorems serves to rule out any positive correlation between judgments or votes — whether due to direct causal dependence or to common cause dependence. Why is it important to rule this out? The following extreme example gives a simple illustration of what happens when the assumption is violated.

**Renowned authority.** A group of three voters is on average moderately competent, with  $c^* = 0.7$ . Were all to vote according to their independent judgments, the probability of a correct majority vote would exceed 0.7 (according to Heterogeneous Competence CJT). However, one of the voters,  $i_j$ , is a renowned authority within the group, yet she is in fact only minimally competent with  $c_j = 0.6$ . Being renowned as she is, her judgment determines all the other voters' judgments. Since they vote according to their authority-determined judgments, the majority of the voters (in fact, all of them) will be no more likely to vote for the correct option than the renowned authority. Thus, the probability of a correct majority vote will be 0.6, lower than average competence. The non-asymptotic conclusion does not hold. Moreover, even if more voters were added to the group, as long as they voted according to their authority-determined judgments, the probability of a correct majority vote would not increase. The asymptotic conclusion does not hold either.

---

<sup>21</sup> There is no need to presuppose a specific theory of causation here. My arguments for the weighted majority rule presuppose only the presence (or absence) of probabilistic correlation between voters or votes. We can then, as here suggested, make sense of such correlation by referring to the causal relations between voters and votes. Yet the question of how to understand the nature of these causal relations is of no direct significance for my arguments.

<sup>22</sup> Boland (1989: 185). Cf. even Grofman, Owen and Feld (1983) and Hawthorne (*mimeo*).

<sup>23</sup> Ladha (1992: 618).

<sup>24</sup> Dietrich and List (2004). Note that the various common causes I list may differ in their effects. E.g. belonging to different ‘schools of thought’ might bring people to judge an issue differently, even though they share the same evidence. Or taking part in mutual deliberation might bring two people to judge similarly, even though each has separate evidence that contradicts the other's.

<sup>25</sup> Cf. Dietrich and Spiekermann (2013). They also have a good discussion of these sources of correlation and of different independence conditions.

The *Renowned authority* case is an extreme illustration of the damaging effect of violating the independence assumption through positive correlation (due to direct causal dependence between voter judgments). Without Independence, Heterogeneous Competence CJT does not hold.

It needs to be pointed out that the independence assumption is not designed to rule out *any* positive correlation between votes (or judgments; I henceforward focus on votes). It does not require *unconditional* independence (unconditionally uncorrelated judgments or votes). In fact, an assumption of unconditional independence would conflict with the relevant competence assumption. Consider a typical Condorcet context, with correctness-tracking voters who vote according to their judgments, and a correct option (according to some independent standard). Suppose that we know that all voters have a competence (probability to judge correctly) of  $c = 0.75$ . Unbeknownst to us, the correct option is  $x$ . Then, when 999 voters have cast their votes, and we have observed a large majority of them to be cast for  $x$ , we can infer that, quite likely,  $x$  is the correct option. Hence, we can infer that, quite likely, the 1.000<sup>th</sup> voter will cast her vote for  $x$  with a 0.75 probability. This means that the votes are probabilistically correlated, in virtue of the assigned competence (correctness-tracking probability). They are not unconditionally independent. Yet they can still be independent in another sense, namely, *state-conditionally* independent. This can be spelled out as follows. Say that we know that  $x$  is the correct option — let us call this the state of the world. Knowing that a large proportion of the 999 votes is cast for  $x$  will tell us nothing new about the probability that the 1.000<sup>th</sup> voter will vote for  $x$ . This is so since, knowing the state — that  $x$  is the correct option — and knowing that everyone has a competence of  $c = 0.75$ , we can infer independently that the 1.000<sup>th</sup> voter will vote for  $x$  with a 0.75 probability. In other words, the state-conditional probability of any voter  $i$ 's vote to be cast for  $x$ , equals her state-conditional probability, given that some other voter  $j$  voted for  $x$ .  $i$ 's and  $j$ 's votes are then probabilistically uncorrelated, conditional on the state.<sup>26</sup>

The independence assumption of *Classical CJT*, as well as the other previously considered Condorcet theorems, presupposes state-conditional independence. This assumption is violated when the votes are probabilistically correlated *given the state*, e.g. when 999 votes still tell us something new about the 1.000<sup>th</sup> voter's probability to vote for  $x$ , even though we know the state of the world and voter competence. As stated, such correlations between votes can be due to, e.g. commonly shared distorting evidence, or the presence of a renowned authority whose vote determines how others vote.

Now, let us return to our here considered context with voters who have unequal stakes in a decision. Consider a case where one voter,  $i$ , has an indi-

---

<sup>26</sup> In a way we might say that the state of the world is a common cause, since it effects correlation between votes (cf. 5.2.8 below). Still, in order to avoid confusion, in the following text, I do not include the state within the class of common causes.

visible bundle of several votes. Then, even if we conditionalise on the state of the world — if we learn that  $i$  casts one of her votes,  $v_1$ , for  $x$ , we will change the probability we assign to another one of her votes,  $v_2$ , of being cast for  $x$ , to certainty.<sup>27</sup> This means that state-conditional independence is violated when we introduce indivisible vote bundles. Our present case is thus in a relevant way analogous to the considered *Renowned authority* case (it would be entirely analogous were  $i$  the only voter), in which the non-asymptotic and asymptotic conclusions no longer hold. All the above arguments for collective optimality rely on the asymptotic conclusion. Without it, the above suggested strategy — to simply and straightforwardly expand this argument, from equal-stakes cases to the domain of unequal-stakes cases — fails.

In the following section, I explore how damaging this violation of the independence assumption is for the case for the weighted majority rule. Note that for now, I assume common-cause independence. That is, I assume that votes are not probabilistically correlated due to any common causes (besides the state of the world, that  $x$  (or  $y$ ) is the common-interest option). However, since common-cause independence — due to, e.g. shared evidence — is still a pretty strong assumption in its own right. I return to it in 5.2.8 and see whether it can be relaxed as well.

### 5.2.7 Unequal-stakes cases and direct causal dependence

We have just noted that direct causal dependence of votes in unequal-stakes cases is relevantly analogous with that in the rather extreme *Renowned authority* case, in which one voter's judgment determines the way all others judge and vote. Let us now consider a small series of examples with dependent voters that are a little less extreme, to get a better understanding of the effects of direct causal dependence between voters. (I get back to dependence between votes at the end of this section.)

To spare myself and the reader an exercise of excessive number-crunching, with large numbers of voters, I first analyse the effects regarding the non-asymptotic conclusion. That is, I consider whether introducing direct causal dependence can have the effect of lowering the probability of a correct majority vote,  $P_n$ , below average voter competence  $c^*$ . In order to do that, I first describe an (equal-stakes) case without direct causal dependence, in which the non-asymptotic conclusion holds and then transform it into a case where there is such dependence due to voting leaders.

***Uncorrelated voters.*** There are five voters in total, with one vote each. We assume their votes to be (state-conditionally) uncorrelated. We can interpret this

---

<sup>27</sup> To be precise, there will be a change in probability only if  $i_j$  is assumed to be less than fully competent.

as there being five voters who vote independently: they have separate evidence, no communication, no ‘schools of thought’, voting leaders, or the like. While voters  $i_1$  and  $i_2$  are highly competent, with  $p_1 = p_2 = 0.9$ , voters  $i_3$ ,  $i_4$  and  $i_5$  are less competent, with  $p_3 = 0.41$  and  $p_4 = p_5 = 0.4$  (see *Table 2*).

	$i_1$	$i_2$	$i_3$	$i_4$	$i_5$
$c_i$	0.9	0.9	0.41	0.4	0.4

*Table 2*

From these figures we can calculate the voters' average competence as  $c^* = 0.602$ . This means that these  $n = 5$  voters' average competence exceeds  $\frac{1}{2} + 1/2n = 0.6$ . We can therefore — according to Distribution-Neutral CJT (see 5.2.5 above) that is applicable to this *uncorrelated* voters case — infer that the probability  $P_n$ , that a majority votes for the correct option, must be greater than the voters' average competence  $c^*$ . That is, in this case, the non-asymptotic conclusion holds.

**Positively correlated voters.** There are again five voters, with the same competences as specified in *Table 2*. However, among the five, there are two ‘voting leaders’ and two ‘followers’, respectively. More specifically,  $i_2$  votes with  $i_1$  (with certainty), while  $i_5$  votes with  $i_4$  (with certainty). In other words, these two pairs are maximally positively correlated. The other eight pairs of voters, we assume, are (state-conditionally) uncorrelated. What are the effects of introducing such positive correlation on the probability that a majority votes for the common-interest option?

Since we know that  $i_2$  votes with  $i_1$ , and  $i_5$  with  $i_4$ , the followers',  $i_2$ 's and  $i_5$ 's, competences levels provide no independent information about their voting behaviour (beyond mirroring their respective voting leader's probability to vote for the correct option). In order to assess the majority outcome we can thus safely ignore the followers' competence levels and represent the case as a three-voter case, according to *Table 3*.

	$i_1$	$i_3$	$i_4$
$c_i$	0.9	0.41	0.4

*Table 3*

Now, we can see that the correct option will be selected under only two specific sets of circumstances: (1) either all three relevant voters vote correctly, or (2) exactly two of them do. If only one (or none) of them votes correctly, the other two (or all three) of these voters constitute a majority against the correct option.

The probability that (1) all three relevant voters vote correctly equals the product of their competences, that is,  $0.9 \cdot 0.41 \cdot 0.4 = 0.1476$ . The probability that (2) exactly two of them vote correctly equals the sum of the products of these two voters' competence levels and the third voter's incompetence level, that is,  $0.1 \cdot 0.41 \cdot 0.4 + 0.9 \cdot 0.59 \cdot 0.4 + 0.9 \cdot 0.41 \cdot 0.6 = 0.0164 + 0.2124 + 0.2214$

= 0.4502. In sum then, the probability that the majority votes correctly equals  $P_n = 0.5978$ . This figure is lower than the voters' average competence,  $c^* = 0.602$ .

This means that, by introducing direct causal dependence between voters, we have transformed a case in which the non-asymptotic conclusion was satisfied into one where it is not. We can now use the *Positively correlated voters* case to illustrate the effects of such dependence on the asymptotic conclusion. Imagine that we start out with a three-voter case, with voters  $i_1$ ,  $i_3$  and  $i_4$ , according to *Table 3*. With only these three voters, the probability of a correct majority vote would be  $P_n = 0.5978$ . Now, we add two voters,  $i_2$  and  $i_5$ , according to *Table 2*, such that  $i_2$  votes with  $i_1$  (with certainty), while  $i_5$  votes with  $i_4$  (with certainty). As we just saw, for these five voters, the probability of a correct majority vote would *not* have increased, but would rather remain exactly the same,  $P_n = 0.5978$ . This suggests that even the asymptotic conclusion is in trouble, when there is direct causal dependence between voters.<sup>28</sup>

It might now be objected that the calculations for the *Positively correlated voters* case are inconsistent. For calculating  $c^*$ , we treated it as a case with five voters, yet for calculating  $P_n$  we treated it as a three-voter case. The objection's upshot is that the non-asymptotic conclusion would in fact hold, if we treated the case as a three-voter case throughout, thereby eliminating the 'followers'  $i_2$  and  $i_5$  from calculations of average voter competence. For a three-voter case according to *Table 3*, average voter competence would be  $c^* = 0.57$ , which means that  $P_n = 0.5978$  would exceed  $c^*$ , just as stated by the non-asymptotic conclusion.

However, this objection is off the mark. We are in fact dealing with a five-voter case, so both  $c^*$  and  $P_n$  should be calculated for the group of five. Yet, when it comes to calculating  $P_n$ , our information about direct causal dependence between voters tells us that certain distributions of votes are rendered impossible — namely, the ones where the respective voting leader and follower vote against each other. This changes the distribution of probabilities *that a majority votes correctly* throughout, such that it resembles the distribution within a three-voter case (as specified in *Table 3*). Yet this does not mean that we are in fact dealing with a three-voter case. There are five voters, and each voter's competence, as the probability *that this individual voter votes correctly*, must be taken into account when calculating the group's average competence  $c^*$ . So we can see that direct causal dependence

---

<sup>28</sup> I formulate this a bit cautiously as a suggestion, since by adding the two 'followers'  $i_2$  and  $i_5$ , the group's average competence rises from  $c^* = 0.57$  to  $c^* = 0.602$ . This violates the condition of 'fixed' average competence, as stated by Boland for Distribution-neutral CJT (see footnote 16 above). It is remarkable, however, that even when average competence is *raised* by adding more voters, the probability of a correct majority vote may *not* increase, when there is direct causal dependence between voters.

between voters may threaten the non-asymptotic conclusion, by lowering the probability that a majority votes correctly.

Interestingly, direct causal dependence can also have the opposite effect when it invokes *negative* correlation between voters: that is, when a pair of judgments or votes is *less* likely to concur than probabilistically expected. To see this, let us return to the above *Positively correlated voters* case and introduce some negative correlation in it.

***Positively and negatively correlated voters.*** There are again five voters with the same competences as specified in *Table 2*. And again, there are two ‘voting leaders’ and two ‘followers’, respectively, such that  $i_2$  votes with  $i_1$  (with certainty), while  $i_5$  votes with  $i_4$  (with certainty). But now, moreover, some of the voters are negatively correlated. More specifically,  $i_3$  is certain to vote against both  $i_4$  and  $i_5$  (thus, we might say that  $i_4$  and  $i_5$  function as ‘obverse voting leaders for  $i_3$ ’). The remaining six pairs of voters, we assume, are (state-conditionally) uncorrelated. What are the effects of introducing negative correlation on the probability of a correct majority vote,  $P_n$ ?

First of all, we cannot assign to these five voters the same competence levels as were specified in *Table 3*. Since  $i_3$  is certain to vote against both  $i_4$  and  $i_5$ ,  $i_3$ 's competence level must equal these two's incompetence levels — that is, we need to raise it from 0.41 to 0.6, according to *Table 4*. This obviously affects average group competence (as compared to the above *Positively correlated voters* case), raising it from 0.602 to  $c^* = 0.64$ . Now, of course, such an increase in individual competence in itself will increase the probability of a correct majority vote — in this case from 0.5978 to  $P_n = 0.708$ .<sup>29</sup>

	$i_1$	$i_2$	$i_3$	$i_4$	$i_5$
$c_i$	0.9	0.9	0.6	0.4	0.4

*Table 4*

Yet note that by introducing negative correlation, the probability that the majority votes for the correct option is actually raised far above that level. Since  $i_3$  with certainty votes against  $i_4$  and  $i_5$ , at least one of these three voters will vote for the correct option. This means that there will be a majority for the correct option whenever  $i_1$  (and with her  $i_2$ ) votes for it, the probability for which is 0.9. Hence, the probability that the majority votes correctly is here in fact  $P_n = 0.9$ .

---

<sup>29</sup> As in the above case, the correct option will be selected under only two specific sets of circumstances: (1) either none of these three relevant voters votes incorrectly, or (2) exactly one of them does. The probability that (1) none of the three relevant voters votes incorrectly equals the product of their competences, that is,  $0.9 \cdot 0.6 \cdot 0.4 = 0.216$ . The probability that (2) exactly one of them votes incorrectly equals the sum of the products of this voters incompetence and the other two's competences, that is,  $0.1 \cdot 0.6 \cdot 0.4 + 0.9 \cdot 0.4 \cdot 0.4 + 0.9 \cdot 0.6 \cdot 0.6 = 0.024 + 0.144 + 0.324 = 0.492$ . In sum then, the probability that the majority votes correctly is  $P_n = 0.708$ .

The just considered three cases show that positive correlation between voters may reduce the probability of a correct majority vote. They also show that negative correlation can increase it, which suggests that the damaging effects of positive correlation — e.g. due to voting leaders — can be mitigated by negative correlation — e.g. when there are ‘obverse’ voting leaders who bring others to vote in opposition to them.

These results are corroborated on a more general level by another Condorcet theorem, which holds for dependent voters. Ladha shows that the asymptotic and non-asymptotic conclusions hold even when votes are correlated — e.g. due to a voting leader — if certain other conditions are satisfied.<sup>30</sup>

Let us call the probability that two different voters  $i$  and  $j$  both vote simultaneously for the correct option *simultaneity-probability*,  $s_{ij}$ .  $s_{ij}$  is defined as a function of  $i$ 's and  $j$ 's competences (their individual probabilities to vote for the correct option) and the coefficient of correlation between their votes. When  $i$ 's and  $j$ 's votes are *uncorrelated*, their simultaneity-probability  $s_{ij}$  equals the product of their competences, that is  $s_{ij} = c_i c_j$ . When  $i$ 's and  $j$ 's votes are *positively correlated*, their probability to simultaneously vote for the correct option is higher than probabilistically expected from their individual competences, that is,  $s_{ij} > c_i c_j$ . When they are maximally positively correlated, such that, e.g.  $j$  is certain to vote with  $i$ ,  $s_{ij}$  equals  $i$ 's competence, that is  $s_{ij} = c_i$ . This is the maximum simultaneity-probability level. Finally, when  $i$  and  $j$  are *negatively correlated*, their probability to simultaneously vote for the correct option is lower than probabilistically expected from their individual competences, that is,  $s_{ij} < c_i c_j$ . And when they are maximally negatively correlated, such that, say,  $j$  is certain to vote against  $i$ , their simultaneity-probability level is at its minimum  $s_{ij} = 0$ .<sup>31</sup>

From the individual simultaneity-probabilities  $s_{ij}$  we can calculate an average  $s^*$  for all pairs of  $n$  voters to vote simultaneously for the correct option. To do that, we sum the  $s_{ij}$  for every pair of voters  $i$  and  $j$  and divide it by the total number of pairs in the group ( $\frac{1}{2}n(n-1)$ ). E.g. for three voters  $i_1$ ,

---

<sup>30</sup> Ladha (1992).

<sup>31</sup> To be precise,  $s_{ij} = \rho_{ij}\sigma_i\sigma_j + c_i c_j$ , where  $\rho_{ij}$  is the coefficient of correlation between  $i$ 's and  $j$ 's votes,  $\sigma_i^2 = c_i(1-c_i)$  is the variance of voter  $i$ 's votes, and  $c_i$  is, as before, voter  $i$ 's competence level (such that  $1-c_i$  is her incompetence level) (cf. Ladha 1992: 625).

This equation can be made intelligible as follows: when  $i$ 's and  $j$ 's votes are *uncorrelated*, the coefficient of their correlation  $\rho_{ij}$  is zero, and so  $s_{ij}$  equals the product of their competences, that is  $s_{ij} = c_i c_j$ . When  $i$ 's and  $j$ 's votes are *positively correlated*,  $\rho_{ij}$  is greater than zero, and so  $s_{ij} > c_i c_j$ . When they are maximally positively correlated, such that, say,  $j$  is certain to vote with  $i$ ,  $\rho_{ij} = 1$ , and  $s_{ij}$  equals  $i$ 's competence, that is  $s_{ij} = c_i$ . This is the maximum simultaneity-probability level. Finally, when  $i$  and  $j$  are *negatively correlated*,  $\rho_{ij}$  is smaller than zero, and  $s_{ij} < c_i c_j$ . When they are maximally negatively correlated, such that, say,  $j$  is certain to vote against  $i$ ,  $\rho_{ij} = -1$ , and  $s_{ij} = 0$ . This is the minimum simultaneity-probability level.

$i_2$ , and  $i_3$ , the average simultaneity-probability  $s^* = (c_{12} + c_{13} + c_{23}) / \frac{1}{2}n(n-1) = (c_{12} + c_{13} + c_{23}) / 3$ . We can now define the following.

**Simultaneity-Probability Condition:** The voters' average simultaneity-probability is smaller than a specific function  $f$  of the numbers of voters  $n$  and their average competence  $c^*$ , such that  $s^* < f(n, c^*)$ , with  $f(n, c^*) = c^{*n} / (n-1) \cdot [(c^* - 1/4)(1 - c^*)] / c^*$ .

With this terminology in place, we can state Ladha's theorem as follows.<sup>32</sup>

**Voter Dependence CJT.** For binary decisions with exactly one correct option (according to some independent standard) and a group of  $n = 2m+1$  voters, given Minimal Average CJT-Competence and Voting According to Judgment, (i') the probability of a correct majority vote  $P_n$  is higher than the voters' average (CJT-)competence  $c^*$ , given the Simultaneity-Probability Condition, and (ii')  $P_n$  strictly increases with the number of voters  $n$  and approaches certainty as the number of voters increases to infinity and as average simultaneity-probability  $s^*$  approaches the squared average competence  $c^{*2}$ .

This theorem differs from all the previous theorems in that it does not rely on an assumption of independence. Yet we should stop to note that it states the non-asymptotic and asymptotic conclusions conditional on a further pair of assumptions. I take a closer look at the non-asymptotic conclusion (i') and its specific Simultaneity-Probability Condition in 5.4 below, where this conclusion is part of an alternative argument for the weighted majority rule. As for now, let us take a look at the asymptotic conclusion (ii'), which is relevant for the present argument. This conclusion states that — given Minimal Average CJT-Competence and Voting According to Judgment — the probability of a correct majority vote approaches certainty as the number of voters increases to infinity *and* as  $s^*$  approaches  $c^{*2}$ . So now, large enough numbers are not quite sufficient; in addition,  $s^*$  must be sufficiently close to the squared average competence  $c^*$ .

To get a more intuitive grasp of what this latter clause means, consider the following simplified case (which is a typical case for Classical CJT). There are  $n$  voters with equal minimal competence  $c$ , and thus average voter competence  $c^* = c$ . Moreover, all voters vote independently. That is, all voter pairs are uncorrelated. This means that for each voter pair  $i$  and  $j$ , their simultaneity-probability  $s_{ij} = c^2$ . Thus, the average simultaneity-probability  $s^* = c^2$ . We can then infer that in such a case with uncorrelated voters,  $s^* = c^{*2}$ . When positive correlation between voters — e.g. due to voting leaders — is introduced in this case,  $s^*$  increases such that  $s^* > c^{*2}$ . The effect of such an increased  $s^*$  on the probability of a correct majority vote,  $P_n$ , has been illustrated by my above *Correlated voters* examples:  $P_n$  in an assumed five-voter case equals  $P_n$  in a three-voter case. This suggests that an *increase*

---

<sup>32</sup> Ladha (1992: 626, 628).

in positive correlation has the same effect as a *decrease* in numbers of voters.<sup>33</sup> And thus, an increasing level of  $s^*$  can counter the effects of increasing numbers  $n$ . Thus, when the numbers  $n$  approach infinity *and* average simultaneity-probability  $s^*$  approaches the level it takes without voter correlation ( $c^{*2}$ ),  $P_n$  increases and approaches certainty. In other words, the closer  $s^*$  is to  $c^{*2}$ , the closer a case with dependent voters resembles a case with independent voters, for which the asymptotic conclusion has been shown to hold.<sup>34</sup>

Now, the discussion in this section has been conducted under the assumption that there is direct causal dependence between *voters* due to there being other voters who function as voting leaders. We may call this form of dependence *interpersonal* direct causal dependence. Yet this study is concerned with dependence between the *votes* held by one and the same voter, whenever she is assigned an indivisible vote bundles. We may thus call this *intrapersonal* direct causal dependence. The previous discussion and the Voter Dependence CJT are of relevance for both forms of dependence. Their effects are the same whether the ‘voting leader’ is another voter or, metaphorically speaking, just another vote within the same bundle of votes. Thus, all the above results can be straightforwardly applied to unequal-stakes cases with indivisible vote bundles. Moreover, this means that the effects of *intrapersonal* direct causal dependence due to indivisible vote bundles, which necessarily invokes positive correlation, may be counterbalanced by *interpersonal* dependence, when this invokes negative correlation between voters.

<sup>33</sup> This suggestion is supported by Hawthorne's (*mimeo*) results. Hawthorne shows that ‘when the average individual competence level is at least slightly better than chance, the more independently people vote, the higher the probability that the majority will select the better policy’ (*mimeo*: 23). Thus, although a high level of independence improves the result, complete independence is not a requirement for the theorem to apply. In fact, Hawthorne observes, for lower degrees of independence, ‘the effect is merely to lower the group competence level to that enjoyed by a somewhat smaller group of more nearly independent voters’ (*mimeo*: 34). For another CJT-extension covering cases with dependent voters, see Boland (1989: 185f.). Boland, however, assumes equal competence and equal voter correlation for all individuals.

<sup>34</sup> Another illustration of the asymptotic conclusion (ii') with its added clause (f) can be provided by my earlier *Uncorrelated*, *Positively correlated* and *Positively and negatively correlated voters* cases. The following table records the relevant variables (where  $P_n$  denotes the probability of a correct majority vote).

	$s^*$	$c^*$	$c^{*2}$	$P_n$
<i>Uncorrelated voters</i>	0.3476	0.602	0.3624	$> c^*$ (according to Distribution-Free CJT)
<i>Positively correlated voters</i>	0.4506	0.602	0.3624	0.5978
<i>Positively and negatively correlated voters</i>	0.452	0.64	0.4096	0.9

Here we can see that  $s^*$  is closest to  $c^{*2}$  in the first, *Uncorrelated voters* case; it is furthest from  $c^{*2}$  in the second, *Positively correlated voters* case; and it is brought somewhat closer again by introducing negative correlation in the third, *Positively and negatively correlated voters* case.

We can then restate the above Voter Dependence CJT in terms of votes rather than voters. That is, we state the theorem for a set of  $n = 2m+1$  votes, which are cast by some number of voters  $l$  (with  $l < n$ ), who are minimally (CJT-)competent on average. For this set of votes we can calculate an average simultaneity-probability  $s^*$ , as the average of the probability that a pair of votes is cast simultaneously for the correct option, for all possible pairs. The thus adapted Vote Dependence CJT can then serve as the first premise in the *fifth argument from collective optimality*.

(1) *Vote Dependence CJT*. For binary decisions with exactly one correct option (according to some independent standard) and set of  $n = 2m+1$  votes cast by any number of voters  $l < n$ , given Minimal Average CJT-Competence and Voting According to Judgment, (i') the probability of a correct majority vote  $P_n$  is higher than the voters' average (CJT-)competence  $c^*$ , given the Simultaneity-Probability Condition, and (ii')  $P_n$  strictly increases with the number of votes  $n$  and approaches certainty as the number of votes increases to infinity and as average simultaneity-probability  $s^*$  approaches the squared average competence  $c^{*2}$ .

(2) The correct option is the common-interest option, according to the independent standard of the given criterion of the common good.

(3) *The Weighted Majority Rule*: For all individuals and any decision with two options, (a) every individual is assigned a number of votes in proportion to her stakes, and (b) the option that receives a majority of votes is selected as outcome.

(4) For any given level of average (CJT-) competence, there is a range of probability levels for a correct majority vote that we call 'near certainty' and a corresponding range of numbers of voters (according to the Condorcet jury theorem) that we call 'sufficiently large numbers' of voters, in combination with a range of average simultaneity-probability  $s^*$  sufficiently close to  $c^{*2}$  that we call 'on average tolerably correlated votes'.

(5) A rule that with near certainty selects the common-interest option, as defined by the given criterion of the common good, is weakly collectively optimal, according to this criterion.

Hence:

(6) For binary decisions with exactly one common-interest option (according to the given criterion of the common good) and set of a sufficiently large number  $n = 2m+1$  of on average tolerably correlated votes cast by any number  $l < n$  voters, given Minimal Average CJT-Competence and Voting According to Judgment, the weighted majority rule is weakly collectively optimal (according to this criterion).

This shows that Condorcet theorems can be applied to the workings of the weighted majority rule, even in cases with indivisible vote bundles due to

unequal stakes, where there is direct causal dependence between votes. If the number of votes is sufficiently large, and they are on average tolerably correlated, the weighted majority rule can be shown to be collectively optimal even with less than full voter competence.<sup>35</sup>

Note that, again, employing Minimal Average CJT-Competence in effect removes the need for any specific assumptions about the voters' desired ends. As long as their individual CJT-competence — as calculated from their individual end-competence — does not bring down average CJT-competence below the given limit of  $\frac{1}{2} + 1/2n$ , it does not matter whether voters are motivated to vote according to common-, self-, or partial-interested ends.

In an unequal-stakes case with many votes tied up in indivisible vote bundles, average simultaneity-probability  $s^*$  may easily be high, and thus not sufficiently close to the squared average competence  $c^*$  to guarantee the weighted majority rule's weak collective optimality. Consider a further twist: a voter's end-competence level and the amount of stakes she holds might be correlated.<sup>36</sup> Plausibly, the more a voter knows to have at stake in a decision, the more motivated she may become to inform herself about the options and their impact on her self-interest and thus the more end-competent she would tend to be. Then, *ceteris paribus*, if voter  $i$  had higher stakes in one decision than in another, and thus a larger vote bundle, there would be a higher level of interdependence in the former than in the latter, with lower group competence as a result. However, if  $i$  in both cases were a majority stake-holder, her higher stakes in the first decision would tend to 'make' her more CJT-competent, such that average group competence might in fact be higher than in the second decision. Thus, the damaging effects of a higher average simultaneity probability  $s^*$  for the group might be mitigated by an increase in average competence  $c^*$ . Unfortunately, if  $i$  instead were a minority stake-holder, her higher stakes in the first decision would tend to 'make' her less CJT-competent (because more end-competent), such that average competence might in fact be much lower than in the second decision. Thus, such a correlation between a voter's stakes and her end-competence might be either beneficial or, really, quite harmful, from a common-good perspective.<sup>37</sup>

In addition, there are other factors that may further increase positive correlation between votes. Imagine, for instance, a group of common-interested voters who all share the same misleading evidence as to what option is in the

---

<sup>35</sup> Someone might wish to suggest an easy way of improving group competence: just give everyone more votes — e.g., multiply every voter's stake-proportional number of votes by 100, or why not 1.000.000. However, this suggestion does not work: all these large vote bundles would increase average vote correlation and hence counteract increasing group competence.

<sup>36</sup> I owe this suggestion to Gustaf Arrhenius.

<sup>37</sup> One could, of course, question the initial assumption that greater stakes tend to motivate voters to get informed, with the Downsian problem of information costs (see Downs 1957).

common interest. The misleading evidence works as a common cause, making it more likely for voters to simultaneously vote *against* the common-interest option. Indeed, such *common-cause dependence* is a concern for many theorists who have worked with Condorcet theorems.<sup>38</sup> Recently, Franz Dietrich and Kai Spiekermann have made an interesting proposal to allow common-cause dependence and formulated a Condorcet theorem that works on a much weaker independence assumption.<sup>39</sup> This is the topic of the next section.

### 5.2.8 Unequal-stakes cases and common cause dependence

Dietrich and Spiekermann take on the problem of positive correlation between voters from a different angle.<sup>40</sup> They consider cases where there is no direct causal dependence (e.g. due to voting leaders), but where there is positive correlation due to *common causes*. Common causes are factors that causally affect more than one voter, such that the affected voters become more likely to simultaneously cast their votes for the same option than they would be in the absence of these factors. In other words, common causes effect positive correlation between votes.

Now, note that the state of the world as described by ‘ $x$  (or  $y$ ) is the correct option’ can be called a common cause for, e.g., all voters who are minimally competent. This is so since the presence of the state makes it more likely that such minimally competent voters simultaneously vote for  $x$ , than

---

<sup>38</sup> See e.g. Rawls (1999: 315), stating that ‘it is [...] clear that the votes of different persons are not independent. Since their views will be influenced by the course of the discussion, the simpler sorts of probabilistic reasoning do not apply’. Cf. Estlund (2008: 225f.) Cf. also Dietrich (2008: 10f.) who argues that when the decision problem faced by the voters is ‘randomly drawn from a reference class of relevant problems’, rather than a specific or ‘fixed’ problem, the independence assumption is ‘usually violated’: there are usually common causes, such as shared evidence, which will lead to correlation between votes. Independence could be vindicated by conditionalising on all common causes, but plausibly common causes will vary over the problems within the considered reference class. In order to conditionalise, we need to focus on a ‘fixed’ decision problem, with its specific common causes. Dietrich also argues, when considering fixed decision problems, we ‘can usually not know [...] whether the voters are competent’, since this requires knowledge of both the correct option and the nature of the specific common causes (2008: 7f.), and this, Dietrich’s contention goes, is something of a tall order. Dietrich’s conclusion is thus the following dilemma: in order to vindicate independence, we usually have to sacrifice our faith in the competence assumption, and in order to vindicate the competence assumption, we usually have to sacrifice our faith in the independence assumption. (Note that the title of Dietrich’s paper puts the point much more strongly; I think it should be modified to state that ‘The premises of Condorcet’s Jury Theorem are not [*Usually*] Simultaneously Justified’.) Dietrich and Spiekerman (2011) offer a way out: they restate the independence and competence assumptions (and hence their entire Condorcet theorem) conditional on the state *in conjunction with all common causes* (see 5.2.8 below).

<sup>39</sup> Dietrich and Spiekermann (2011).

<sup>40</sup> Dietrich and Spiekermann (2011).

its absence (as described by ‘there is no correct option’). However, correlation due to the state is not a problem in the light of the CJT-independence assumption. What the latter requires is not unconditional independence between votes, but *state*-conditional independence. I therefore here use the term ‘common cause’ for all factors that effect positive correlation between votes *except* the state of the world as described by ‘ $x$  (or  $y$ ) is the correct option’.<sup>41</sup>

Dietrich and Spiekermann's strategy to deal with positive correlation differs from Ladha's (see 5.2.7 above). Recall that Ladha restates the Condorcet theorem such that it allows a certain degree of dependence between voters.<sup>42</sup> Dietrich and Spiekermann instead set out to treat *Correlated voters* cases as cases in which voters nevertheless are independent — in another sense of the word. They do this by conditionalising the voters' probability to vote for the correct option not only on the state, but also on all common causes. The conjunction of all common causes and the state of the world that voters vote on is called ‘the problem’,  $\pi$ . Dietrich and Spiekermann then state that ‘by conditionalizing on the problem, we fix all those circumstances which, if left variable, could lead to voter correlation’.<sup>43</sup> Since they assume that there is no correlation due to direct causal dependence, voters are thus assumed to vote uncorrelatedly, *given* the problem  $\pi$ . I say that the voters vote ‘*problem-conditionally independent*’.<sup>44</sup> Let me illustrate this with a simple example.<sup>45</sup>

**Shared evidence.** There are three voters,  $i_1$ ,  $i_2$  and  $i_3$ , who face a binary decision. We know that they are all equally minimally competent, say,  $c_i = 0.6$ . If they vote state-conditionally independent, then if we know the state (say, that  $x$  is the correct option), we cannot infer a change in  $i_1$ 's probability to vote for  $x$  from the observation that  $i_2$  and  $i_3$  in fact vote for  $x$ .

However, assume that in this case we also know that they all share the same evidence. This evidence is either very good or rather poor — both possibilities are equally likely. If the evidence is good, this means that the three voters are dealing with an *easy problem* (consisting of the state in conjunction with, in this case, clear and reliable evidence for this state). Say that with such good evidence every voter is 0.9 certain to vote correctly. If, however, the evidence is poor, this means that they are dealing with a *difficult problem* (consisting of the state in conjunction with, in this case, complex or misleading evidence concerning this state). Say that with such poor evidence each is only 0.3 certain to vote correctly. Thus ‘overall’ competence for each voter is  $c_i = \frac{1}{2} 0.9 + \frac{1}{2} 0.3 = 0.6$ , as assumed above.

But then, if we observe that  $i_2$  and  $i_3$  in fact vote for  $x$  (rather than that one or both of them vote against  $x$ ), we can infer that, most likely, the evidence is

---

<sup>41</sup> Cf. footnote 25 above.

<sup>42</sup> Ladha (1992).

<sup>43</sup> Dietrich and Spiekermann (2011: 10).

<sup>44</sup> For this term, cf. Dietrich and Spiekermann (2010: 9).

<sup>45</sup> The example and exposition of how independence is violated and restored in another meaning are my own. For a case with a similar structure, see Dietrich and Spiekermann (2011: 11).

good, and that hence  $i_i$  has a 0.9 chance of voting correctly as well. This means that their votes are not state-conditionally independent here. Only if we conditionalise on the problem (the conjunction of the state and the very good evidence) are votes once again independent. Then, knowing how the other two vote will not lead to a change in the probability to vote correctly which we assign to the third voter.

So independence can be restored in another meaning — as problem-conditional independence — by conditionalising on the problem  $\pi$ . Let us therefore define:

***Problem-Conditional Independence:*** Every voter judges (probabilistically) independent *given* the problem  $\pi$  (the conjunction of all common causes and the state of the world that voters vote on).

However, it does not work to simply replace the (state-conditional) Voter Independence assumption by Problem-Conditional Independence and leave the other parts of Classical CJT intact. Consider what the setting in the above *Shared evidence* case implies for the probability of a correct majority decision. If the voters are facing an easy problem, such that all three are 0.9 likely to vote correctly, the probability of a correct majority decision  $P_n(\text{easy}) = 0.972$ .<sup>46</sup> But if they are facing a difficult problem, such that all three are only 0.3 likely to vote correctly, the probability of a correct majority decision  $P_n(\text{difficult}) = 0.216$ .<sup>47</sup> Since both scenarios are equally likely, the overall probability of a correct majority decision is  $P_n = \frac{1}{2} 0.972 + \frac{1}{2} 0.216 = 0.594$ . This is lower than individual ('overall') competence  $c_i = 0.6$ . Hence Classical CJT's non-asymptotic conclusion is violated.

Some further reflection reveals that the asymptotic conclusion is also violated. Greater numbers of voters would increase  $P_n$  (easy) and have it approach certainty as numbers approach infinity. Yet greater numbers would also decrease  $P_n$  (difficult) and have it approach zero, as numbers approach infinity. (This is the 'sad flipside' result of *Classical CJT*, as mentioned in 5.2.1 above.) Thus, the overall probability of a correct majority decision,  $P_n$ , would approach  $\frac{1}{2}$  as numbers approach infinity.

The trouble is that we have restated the independence assumption in problem-conditional terms but retained a problem-insensitive ('overall') competence assumption. Dietrich and Spiekermann provide a remedy by reformulating the latter assumption in terms of an individual voter  $i$ 's competence *regarding a specific problem*  $\pi$ :  $c_i^\pi$ : 'Problem-specific competence [ $c_i^\pi$ ] is more likely to be high than low [and] the same for all voters  $i$ '.<sup>48</sup> That is,

<sup>46</sup> There is a majority for the correct option if and only if all three, or exactly two, of the voters vote correctly, that is:  $0.9^3 + 3(0.1 \cdot 0.9^2) = 0.972$ .

<sup>47</sup> There is a majority for the correct option if and only if all three, or exactly two, of the voters vote correctly, that is:  $0.3^3 + 3(0.7 \cdot 0.3^2) = 0.216$ .

<sup>48</sup> Dietrich and Spiekermann (2011: 12f.).

voters are assumed to be more likely to face easy problems, where every voter  $i$  has the same  $c_i^\pi > 1/2$ , than difficult problems, where every voter  $i$  has the same  $c_i^\pi < 1/2$ . So rather than assuming that voters are better than chance overall, over a sequence of decisions, it is assumed that they are more likely to face decisions on which they are better than chance, than decisions on which they do worse than chance. We can thus define:

**Problem-Conditional Competence:** All voters are more likely to face easy problems, where every voter  $i$  has the same problem-specific competence  $c_i^\pi > 1/2$ , than difficult problems, where every voter  $i$  has the same  $c_i^\pi < 1/2$ .

Dietrich and Spiekermann then propose a new extension of the Condorcet jury theorem.<sup>49</sup> I call it Problem-Specific CJT and reconstruct it as follows (in terms of votes rather than voters).

**Problem-specific CJT:** For binary decisions where exactly one option is correct, and a given number of  $n = 2m+1$  votes, given Problem-Conditional Competence, Voting According to Judgment, and Problem-Conditional Independence, (i'') the probability of a correct majority vote,  $P_n$ , is greater than the voters' problem-specific competence  $c_i^\pi$ ,<sup>50</sup> and (ii'')  $P_n$  approaches certainty if (a) the probability of facing an easy problem is 1 — or approaches a lower limit if (b) the probability of facing an easy problem is lower than 1 — as the number of votes increases to infinity.<sup>51</sup>

As we can see, this version of the asymptotic conclusion differs from all the previous ones. It retains the previous optimistic result of nearly infallible large-number groups, but only on the condition that there is no risk for the voters to face a non-easy problem. If there is the slightest such risk, the result is the much more modest one of clearly fallible large-number groups — which, though, perform better than smaller groups. On the other hand, the non-asymptotic conclusion is preserved — though in terms of problem-specific competence  $c_i^\pi$  rather than average voter competence  $c^*$  — in the face of common-cause dependence.

---

<sup>49</sup> Dietrich and Spiekermann (2011: 12f.).

<sup>50</sup> Dietrich and Spiekermann (2011: 14) non-asymptotic conclusion differs slightly from mine; they state: ‘As the group size increases, the probability that a majority votes correctly (i) increases’. Still, this implies that, as group size  $n$  increases (for any  $n = 2m+1$ ) from one voter  $i$  to more than one voter,  $P_n$  increases from  $P_n = p_i^\pi$  (in the one-voter case, the probability of a correct majority vote equals this one voter's competence) and is thus greater than  $p_i^\pi$ . Thus, Dietrich and Spiekermann's non-asymptotic conclusion implies my non-asymptotic conclusion.

<sup>51</sup> To be more precise, Dietrich and Spiekermann (2011: 14) state that ‘the value to which the probability of a correct majority converges is [...] the probability that the problem is easy [i.e. that  $p_i^\pi > 1/2$ ] plus half of the probability that the problem is on the boundary between easy and difficult [i.e., that  $p_i^\pi = 1/2$ ]’.

There are two problems for employing Problem-Specific CJT to make a case for the weighted majority rule. First, the theorem assumes that voters are independent conditional on the problem. In unequal stakes-cases, with indivisible vote bundles, there is positive correlation between votes even when we have conditionalised on the state and all common causes. As Dietrich and Spiekermann explicitly point out, direct causal dependence between voters violates the assumption of problem-conditional independence.<sup>52</sup> This means that Problem-Conditional CJT cannot be applied to my cases. It might be possible to combine Problem-Conditional CJT and Vote Dependence CJT, to create a theorem that can deal with correlation between votes due to all kinds of causal connections — direct causal dependence between votes as well as common-cause dependence. However, in the absence of such a theorem and its proof, the best we can conclude is this. The argument from weak collective optimality can be stated for unequal-stakes cases without common-cause dependence — on the basis of Vote Dependence CJT. It could also be adapted to cases with common-cause dependence without any form of direct causal dependence (e.g. equal-stakes cases) — on the basis of Problem-Conditional CJT. But as of now, it cannot be stated for the most interesting cases with unequal stakes *and* common-cause dependence.

The second problem with Problem-Specific CJT is that it assumes *homogeneous* problem-specific competence  $c_i^\pi$  for all voters. For self- and common-interested voters with different ends, this assumption has some implausible implications. In the context of this study, it means that we have to assume an equal problem-specific *CJT*-competence for all voters. This implies, if we assume a better-than-chance problem-specific *CJT*-competence for all voters, that all the minority stake-holders' problem-specific *end*-competence must be equally worse than chance, and all the majority stake-holders' problem-specific *end*-competence equally better than chance (by the same margin). In other words, all minority stake-holders must be assumed to be equally end-incompetent, and all majority stake-holders to be equally minimally end-competent. (The reverse holds for an assumed worse-than-chance problem-specific *CJT*-competence.) This is a rather artificial assumption, which threatens to weaken the case for the weighted majority rule.<sup>53</sup> It therefore seems to me that a stronger argument from weak-collective optimality is pending on the development of an extension of Problem-Conditional CJT, which allows *heterogeneous* problem-specific competence.

This section must thus conclude by pointing out the need for further development of the Condorcet theorems, in order to be fully applicable to the present case for the weighted majority rule.

---

<sup>52</sup> Dietrich and Spiekermann (2011: 21).

<sup>53</sup> Cf. footnote 9 above, where I criticise a similar assumption as overly artificial.

### 5.3 Conclusions

The arguments in this chapter employed a number of Condorcet theorems to build a new case for the weighted majority rule, for binary decisions with less than fully competent voters. From the outset, voters were assumed to have their self-interest or the common interest as their desired end, to have correct beliefs regarding which option promotes their respective end, and to vote accordingly. The strong assumption of full voter competence was then relaxed to allow certain better-than-chance competence levels.

Note that the results I derived in this chapter apply to odd-numbered groups of voters. As stated in 5.2.1 above, these results can be extended to even-numbered groups.<sup>54</sup> I state the qualifying term ‘odd’ in the summary of my results below in parenthesis to account for its being inessential to the argument.

For equal-stakes cases with common-interested voters, the *first argument from weak collective optimality* established that the weighted majority rule selects the common-interest option with near certainty when the strong assumption of full voter competence is replaced by an assumption of equal minimal (CJT-)competence and, in addition, (state-conditional) independence is assumed between voter judgments, given a sufficiently large (odd) number of voters. For equal-stakes cases with self-interested voters, the *second argument* established the weighted majority rule's weak collective optimality conditional on the assumptions of equal minimal *end*-competence and independence, again given a sufficiently large (odd) number of voters. The *third argument* showed, by a series of steps from *Only self-interested* to *Only common-interested voters* cases that the same conclusion holds for equal-stakes mixed-motivation cases. The *fourth argument* allowed us to relax the assumption of equal minimal end-competence, by showing that the weak collective optimality of the weighted majority rule is preserved as long as the voters' end-competences result in an average CJT-competence of at least  $\frac{1}{2} + \frac{1}{2n}$ , given independence and sufficiently large (odd) numbers of voters. Formulating the competence condition in terms of average CJT-competence in effect removes the need for any specific motivating assumption: voters may vote with any common-, self- or partial-interested desired end, as long as their individual CJT-competence (as calculated from their end-competence) does not bring down average CJT-competence below the given limit of  $\frac{1}{2} + \frac{1}{2n}$ .

For unequal-stakes cases with self- or common-interested voters, I showed that the assumption of independence is violated, since votes within indivisible vote bundles are positively correlated. However, the *fifth argument* showed that the weighted majority rule is weakly collectively optimal, conditional on minimal average CJT-competence, given a sufficiently large

---

<sup>54</sup> See footnote 3 above for a precisification of this claim.

(odd) number of on average tolerably correlated votes. Again, formulating the condition in terms of average CJT-competence removes the need for any specific motivating assumption: voters may be common-, self- or partial-interested, as long as their individual end-competences result in a sufficiently high average CJT-competence.

A potential further argument from weak collective optimality, which would allow not only correlation due to direct causal dependence within vote bundles, but could moreover even handle additional common-cause dependence between voters by conditionalising on it, has been pointed out to be pending on further development of Condorcet theorems.

Conducting the argument in these steps has had the advantage that we could accurately see which prices we have to pay to adapt the argument from weak collective optimality to different contexts — from equal-stakes common-interested cases to unequal-stakes cases with mixed motivation. For all of these cases, full voter competence, concerning which option is in accordance with the voters' respective desired end, is not required for a compelling case for the weighted majority rule. Overall, the above strong Competence assumption can be relaxed, at the cost of settling for the weak (rather than strong) collective optimality of the weighted majority rule.

However, this result depends heavily on the condition that there are sufficiently many voters (or votes). This means that within the traditional political domain of elections and referenda, with large electorates, the weighted majority rule emerges as a promising candidate from a common-good perspective. Still, for small-scale collective decision-making, weak collective optimality could be preserved only for much higher CJT-competence levels. In the next chapter, I use the Condorcet theorems of this chapter to construct an argument for small numbers of voters (or votes).

To conclude, my arguments in this chapter show that Condorcet theorems are applicable not only within judgment-aggregative accounts of democracy, but can be a useful tool even within other approaches. The classical Condorcet jury theorem, as Ladha puts it, 'assumes that the members of a group who choose between a pair of alternatives (a) share a common goal, (b) vote (statistically) independently, and (c) vote for the better alternative with a probability greater than 0.5'.<sup>55</sup> Recent literature in this field has developed our understanding as to how far assumptions (b) and (c) can be relaxed in order for some version of the Condorcet-results to hold.<sup>56</sup> I have applied

---

<sup>55</sup> Ladha (1993: 69).

<sup>56</sup> There has also been some development concerning a fourth implicit assumption not mentioned by Ladha (1993): (d) voters vote sincerely (or non-strategically). This assumption has been challenged within a game-theoretic approach to voting, in which the individual voter's probability to vote correctly is not understood as her individual competence but rather as the other voters' confidence in the correctness of her vote. For one of the pioneering papers of this approach, see Austen-Smith and Banks (1996). Cf. even Dietrich (2008). It is an interesting further question — albeit one that goes beyond the limits of the present study — whether and

these latter results within the context of my study. Yet here, moreover, even assumption (a) has been relaxed: voters need not be assumed to share the common goal of the common good, but may have their own self-interest (or, indeed, any interest or cause whatsoever) as their goal — as long as their competence regarding this goal translates into required levels of (average) CJT-competence.<sup>57</sup> In addition, I have shown that this result holds for a novel democratic decision rule, whose (weak) collective optimality heretofore is largely unexplored. (As a corollary, since the weighted majority rule is extensionally equivalent to the simple majority rule in equal-stakes cases, my results apply to the latter as well.)

The present work thus aims to open up a new field of investigation, in which the many results of the rapidly growing literature on the Condorcet jury theorem might lead to further interesting insights on non-judgment aggregative voting in general, and the weighted majority rule in particular. For instance, one might wonder how common-cause dependence (in addition to direct causal dependence) for heterogeneous competence levels (rather than equal competence levels) might affect the results. Such questions go beyond the present investigation, but are worth taking up another day.

---

how the case for the weighted majority rule could be restated within such a game-theoretic model.

<sup>57</sup> There have been other such efforts; see e.g. Miller (1986).

## 6 Two more arguments and one alleged paradox

### 6.1 Introduction

In this chapter, I draw on the results of the previous one and use them to devise two further arguments for the weighted majority rule. In 6.2 I restate the argument from weak collective optimality in terms of voting behaviour, rather than in terms of the desires and beliefs (or judgments) of the voters. This allows us to make a case for the weighted majority rule without assuming the voters to display any specific mental structure. In 6.3 I explore how the non-asymptotic conclusions of the considered Condorcet theorems can be employed in an argument for the weighted majority rule that does not rely on large numbers of voters. The argument asserts that this rule is better than the average voter's vote at selecting the common-interest option. In 6.4 I discuss the so-called discursive paradox (or discursive dilemma). This discussion shows that the way the options are framed — how they are set up on the agenda — may have repercussions on voter competence and hence, as an upshot, on the weak collective optimality of the weighted majority rule. Section 6.5 concludes.

### 6.2 The behavioural argument from weak collective optimality

Throughout the last chapter, I have formulated the different Condorcet theorems in terms of the voters' competence. Voters are assumed to be not fully, but sufficiently competent at judging their self-interest or the common interest. Each of the various suggested competence assumptions must be conjoined with further assumptions in order to imply that voters cast their votes as they assumedly do. I proposed, in addition to the competence assumptions, that the voters are assumed to have their self-interest, or the common interest, as a desired end, and to vote according to their judgments made in the light of their ends (or, in another formulation, be free from internal or external obstacles for voting in accordance with their belief-desire pair). I

made, in short, a number of assumptions concerning the voters' motivational set-up.

These assumptions are rather specific. They assign a lot of mental structure (plus other things) to voters. Maybe this is a bit too strong. And obviously, these assumptions are merely jointly sufficient — yet neither jointly nor separately necessary — to establish that voters (can be expected to) vote as required for my arguments for the weighted majority rule. Consider that a voter might be partial-interested (e.g. desiring to promote her family's collective interest), but mistakenly judge that the option that is in her self-interest is the partial-interest option she seeks. Given the Voting According to Judgment (or Success) assumption, she can then be assumed to vote non-erratically (for her self-interest option) after all. Countless other such scenarios can be aligned — all showing that the voting behaviour required for establishing the weighted majority rule's weak collective optimality is not limited to cases in which these specific assumptions hold.

Therefore, it is worthwhile to restate the last chapter's results in terms of voting behaviour alone — regardless of the assumed motivational set-up behind it. More specifically, we can restate Classical CJT, as well as all the other theorems, in purely behavioural terms, that is, in terms of the voters  $i$ 's *reliability*  $r_i$  of voting correctly — rather than in terms of their *competence*  $c_i$  to correctly judge the options (and then, given the other assumptions, to vote accordingly). We can then apply these theorems as premises in the different arguments from weak-collective optimality, which focus exclusively on voter behaviour, without relying on specific motivational assumptions.

To constitute just one example, let us state Vote Dependence CJT (see 5.2.7 above) in purely behavioural terms. Since this theorem is formulated in terms of votes (rather than in terms of voters who cast these votes), we can state it for the votes  $v$ 's *reliability*  $r_v$ .  $r_v$  is simply  $v$ 's probability of being cast for the correct (common-interest) option. The *average* reliability of all the votes,  $r^*$ , is the sum of their individual reliabilities  $r_v$  divided by the number of votes  $n$ . Average reliability of all votes is *minimal* if and only if it is above chance. We can then define:

***Minimal Average Reliability:*** The votes are minimally reliable on average, that is, the average of their individual probabilities to be cast for the correct option (according to the independent standard) is above chance,  $r^* > \frac{1}{2}$ .

Moreover, the simultaneity-probability for any pair of votes is the probability that they are cast simultaneously for the correct option. If a pair of votes  $v$  and  $w$  is independent, its simultaneity-probability  $s_{vw}$  equals the product of the individual votes' reliabilities  $r_v$  and  $r_w$ . If they are positively correlated,  $s_{vw}$  is higher than that; if they are negatively correlated, it is lower (all according to what was said in 5.2.7 above, for correlation between pairs of voters). We can now define the following.

**Simultaneity-Probability Condition (For Votes):** The votes' average simultaneity-probability is smaller than a specific function  $f$  of the numbers of votes  $n$  and their average reliability  $r^*$ , such that  $s^* < f(n, r^*)$ , with  $f(n, r^*) = r^{*n} / (n-1) \cdot [(r^*-1/4)(1-r^*)] / r^*$ .

Now, we adapt Vote Dependence CJT as follows, by simply replacing average voter competence  $c^*$  with average vote reliability  $r^*$ .

**Behavioural CJT:** For binary decisions with exactly one correct option (according to some independent standard) and a set of  $n = 2m+1$  votes cast by any number of voters  $l < n$ , given Minimal Average Reliability, (i') the probability of a correct majority vote  $P_n$  is higher than the votes' average minimal reliability  $r^*$ , given Simultaneity-Probability Condition (For Votes), and (ii')  $P_n$  strictly increases with the number of votes  $n$  and approaches certainty as  $n$  increases to infinity and as average simultaneity-probability  $s^*$  approaches the squared average reliability  $r^{*2}$ .

This theorem can then be employed as the first premise in the *behavioural argument from weak collective optimality*.

(1) *Behavioural CJT*.

(2) The correct option is the common-interest option, according to the independent standard of the given criterion of the common good.

(3) *The Weighted Majority Rule:* For all individuals and any decision with two options, (a) every individual is assigned a number of votes in proportion to her stakes, and (b) the option that receives a majority of votes is selected as outcome.

(4) For any given level of average (CJT-)competence, there is a range of probability levels for a correct majority vote that we call 'near certainty' and a corresponding range of numbers of voters (according to the Condorcet jury theorem) that we call 'sufficiently large numbers' of voters, in combination with a range of average simultaneity-probability  $s^*$  sufficiently close to  $c^{*2}$  that we call 'on average tolerably correlated votes'.

(5) A rule that with near certainty selects the common-interest option, as defined by the given criterion of the common good, is weakly collectively optimal, according to this criterion.

Hence:

(6) For binary decisions with exactly one common-interest option (according to the given criterion of the common good) and set of a sufficiently large number  $n = 2m+1$  of on average tolerably correlated votes cast by any number of voters  $l < n$ , given Minimal Average Reliability, the weighted majority rule is weakly collectively optimal (according to this criterion).

Thus, the weak collective optimality of the weighted majority rule can be established for all binary cases where votes are on average more likely than not to be cast for the common-interest option — no matter how and why this might be the case — as long as there is a sufficiently large odd number of on average tolerably correlated votes.

### 6.3 The better-than-the-average-voter argument

We have in the previous and in the present chapter considered a number of arguments that show the weighted majority rule to be weakly collectively optimal — for sufficiently large groups of (on average tolerably correlated) voters (or votes). Yet for collective decision-making within very small groups (with very few votes), weak collective optimality could be preserved only for very high CJT-competence levels. However, for these cases an alternative argument for the weighted majority rule could be suggested, drawing on the non-asymptotic conclusions of the Condorcet theorems that were stated in the previous chapter. These non-asymptotic conclusions had no relevance for the arguments from weak collective optimality, but can be usefully employed in an argument that works even in the absence of large numbers.

In this section, I propose an argument that takes as its first premise one of the above Condorcet theorems, namely Vote Dependence CJT (see 5.2.7 above). It should be noted that the argument could be adapted to the other stated Condorcet theorems as well (e.g. as a ‘better than any single voter’ argument, with Classical CJT as its first premise). Let us first recapitulate the theorem (omitting the now irrelevant asymptotic conclusion).

***Vote Dependence CJT:*** For binary decisions with exactly one correct option (according to some independent standard) and set of  $n = 2m+1$  votes cast by any number of voters  $l < n$ , given Minimal Average CJT-Competence and Voting According to Judgment, (i') the probability of a correct majority vote  $P_n$  is higher than the voters' average (CJT-)competence  $c^*$ , given the Simultaneity-Probability Condition.

Let us also recall the definitions of the stated assumptions.

***Minimal Average CJT-Competence:*** The voters are minimally (CJT-)competent on average, that is, the average of their individual probabilities to judge the options correctly (according to the independent standard) is above chance,  $c^* > 1/2$ .

***Voting According to Judgment:*** Every voter votes according to this judgment.

**Simultaneity-Probability Condition:** The voters' average simultaneity-probability is smaller than a specific function  $f$  of the numbers of voters  $n$  and their average competence  $c^*$ , such that  $s^* < f(n, c^*)$ , with  $f(n, c^*) = c^* - n/(n-1) \cdot [(c^* - 1/4)(1 - c^*)]/c^*$ .

Now, let us take a look at the theorem's non-asymptotic conclusion (i'). It states that the probability of a correct majority vote is better than average voter (CJT-)competence — given Minimal Average CJT-Competence and Voting According to Judgment — if the following inequality holds:

$$s^* < f(n, c^*), \text{ that is,}$$

$$s^* < c^* - n/(n-1) \cdot [(c^* - 1/4)(1 - c^*)]/c^*.$$

What are we to make of this? Regarding the right-hand side of the inequality:  $f(n, c^*)$  is simply a function of the number of votes  $n$  and the voters' average competence  $c^*$ . This function gives us an upper limit below which the average simultaneity-probability  $s^*$  is sufficiently low to guarantee the non-asymptotic conclusion.<sup>58</sup> Then, we can read off the following: as the average voter competence  $c^*$  increases toward 1, this upper limit increases toward  $c^*$ . This is so since as  $c^*$  approaches 1,  $1 - c^*$  approaches 0, so the right-hand side approaches  $c^*$ , and thus its maximum value. On the other hand, as the average voter competence  $c^*$  decreases toward 0.5, the upper limit decreases toward  $0.5 - [n/(n-1)] \cdot 0.25$  — which, for very large numbers of  $n$ , approaches 0.25 as its minimum value.<sup>59</sup>

Regarding the left-hand side of the inequality: recall that  $s^*$ , the average probability for all pairs of votes to be cast simultaneously for the correct option, is the sum of  $s_{vw}$  for every pair of votes  $v$  and  $w$ , divided by the total

<sup>58</sup> Note that the Simultaneity-Probability Condition is sufficient, but not necessary for (i') to hold. That is, there are cases where it does not hold, but the non-asymptotic conclusion is satisfied nonetheless. My above *Uncorrelated voters* and *Positively and negatively correlated voters* scenarios are cases in point. The following table records the relevant variables (where  $P_n$  denotes the probability of a correct majority vote).

	$s^*$	$f(n, c^*)$	$c^*$	$P_n$
<i>Uncorrelated voters</i>	0.3476	0.3111	0.602	$> c^*$ (according to <i>Distribution-free CJT</i> )
<i>Positively correlated voters</i>	0.4506	0.3111	0.602	0.5978
<i>Positively and negatively correlated voters</i>	0.452	0.3658	0.64	0.9

Also note that the Simultaneity-Probability Condition now provides us with a general guideline on how  $P_n$  might be raised above  $c^*$  in the *Positively correlated voters* case; namely through negative correlation between pairs of voters, which would lower  $c^*$  below  $f(n, p^*)$ .

<sup>59</sup> We can also note that the upper limit  $f(n, c^*)$  decreases slightly with rising levels of  $n$ . This is so since, as  $n$  increases to infinity,  $n/(n-1)$  approaches 1,  $n/(n-1) \cdot [(c^* - 1/4)(1 - c^*)]/c^*$  gets larger, such that  $f(n, c^*)$  decreases toward  $c^* - [(c^* - 1/4)(1 - c^*)]/c^*$ . This decrease is, however, not very significant and gets even less so for very large numbers of  $n$ . Therefore, I do not consider it further.

number of pairs in the group. Compared to a group of uncorrelated votes, the more positively correlated vote pairs there are, and the higher their degree of positive correlation, the higher their individual simultaneity-probabilities  $s_{vw}$ , and thus the higher the level of the average,  $s^*$ . And the higher  $s^*$ , the more likely that it will exceed the upper limit  $f(n, c^*)$ , which might mean that the non-asymptotic conclusion is jeopardised.<sup>60</sup> However, the more negatively correlated vote pairs there are, and the higher their degree of negative correlation, the lower their individual simultaneity-probabilities  $s_{vw}$ , and thus the lower the level of the average,  $s^*$ . And the lower  $s^*$ , the more likely that it will be below the upper limit  $f(n, c^*)$ . Being below the upper limit *guarantees* that the non-asymptotic conclusion holds. In other words, the non-asymptotic conclusion of Vote Dependence CJT provides us with a formula that specifies how much positive correlation between pairs of votes is tolerable, and moreover how negative correlation can mitigate the damaging effects of positive correlation.

Vote Dependence CJT can now be employed as the first premise in yet another argument for the weighted majority rule. I call it the *better-than-the-average-voter argument*.

(1) *Vote Dependence CJT*.

(2) The correct option is the common-interest option, according to the independent standard of the given criterion of the common good.

(3) *The Weighted Majority Rule*: For all individuals and any decision with two options, (a) every individual is assigned a number of votes in proportion to her stakes, and (b) the option that receives a majority of votes is selected as outcome.

(4) A rule that selects the common-interest option with a probability that is greater than the voters' average (CJT-)competence performs better than the average voter, according to the given criterion.

Hence:

(5) For binary decisions with exactly one correct option (according to the given criterion of the common good) and set of  $n = 2m+1$  votes cast by any number of voters  $l < n$ , given Minimal Average CJT-Competence, Voting According to Judgment, and the Simultaneity-Probability Condition, the weighted majority rule performs better than the average voter (according to the given criterion).

---

<sup>60</sup> I say that the non-asymptotic conclusion *might* be jeopardised, since the Simultaneity-Probability Condition is a sufficient, not a necessary condition. However, at one point or another,  $s^*$  will be sufficiently high to undermine the non-asymptotic conclusion. Consider the case where all voters are maximally positively correlated — they all vote as one. Then, the probability of a correct majority vote will not exceed their average competence  $c^*$ . Instead, it will equal  $c^*$ .

This argument states the conditions under which the weighted majority rule is more likely than the vote of the average voter to select the common-interest option.<sup>61</sup> This conclusion can be interpreted as follows. The argument states the conditions under which the weighted majority rule is a better instrument (in the light of the relevant criterion) for selecting the common-interest option than simply relying on a random voter's vote.

However, we should note that a *specific single* voter might of course perform much better than the average, and also better than the weighted majority rule, when voting. It might even be the case that a *random* voter would perform better, when asked not to cast a vote according to whatever motivates her, but rather to directly (and honestly) identify the common-interest option. Further, in unequal-stakes cases with indivisible vote bundles (and, possibly, even common causes), average simultaneity-probability may easily be too high for the conclusion to apply at all. Moreover, outperforming the average voter may not be deemed a very promising feature of the weighted majority rule. To say the least, this is not a plausible interpretation of 'being collectively optimal'. Moreover, taking the vote of the average voter (that is, of a randomly chosen voter) as decisive seems as an unpromising alternative for collective decision-making anyway.<sup>62</sup>

In other words, we should note that the *better-than-the-average-voter argument* not only relies on rather specific and demanding assumptions, but also gives us a somewhat unpromising conclusion. I therefore refrain from developing this line of argument any further.

## 6.4 Some remarks on the discursive paradox

Consider the following case.

**Discursive paradox.** There are three voters,  $i_1$ ,  $i_2$  and  $i_3$  who face two decisions: whether proposition  $a$  is true or false, and whether proposition  $b$  is true or false. While  $i_1$  and  $i_2$  consider  $a$  to be true,  $i_3$  considers it to be false. And while  $i_2$  and  $i_3$  consider  $b$  true,  $i_1$  considers it false. They vote — say, for simplicity — with one vote each. Then, by simple majority, both  $a$  and  $b$  are collectively judged to be true.

However, imagine now that these three voters face the following decision: whether  $a\&b$ , the conjunction of  $a$  and  $b$ , is true or false. As stated,  $i_2$  considers both  $a$  and  $b$  to be true, so she consistently judges their conjunction to be true and votes accordingly.  $i_1$  and  $i_3$ , on the other hand, consider one of the

---

<sup>61</sup> We also need to keep in mind that the argument is stated for odd-numbered groups of voters. As observed in footnote 3 above, Classical CJT's non-asymptotic conclusion can be derived even for even-numbered cases, but this requires a stricter competence assumption (the more strict the smaller the group). The same arguably holds for Vote dependence CJT.

<sup>62</sup> Note that even the voting lottery has its advocates — though from a criterion of fairness rather than of an aggregative common good. See e.g. Saunders (2008).

separate propositions  $a$  and  $b$  to false, so they consistently judge their conjunction to be false and vote accordingly. Now, by simple majority,  $a\&b$  is collectively judged to be false. This means that, from coherent sets of individual judgments on a pair of propositions  $a$  and  $b$  and their conjunction, by simple majority rule we derive an incoherent set of collective judgments: that both  $a$  and  $b$  are true, yet that their conjunction is false (see *Table 1*).<sup>63</sup>

	$i_1$	$i_2$	$i_3$	<i>simple majority outcome</i>
$a$	true	true	false	true
$b$	false	true	true	true
$a\&b$	false	true	false	false

*Table 1*

So-called discursive paradoxes (or discursive dilemmas), such as the above, are discussed within an emerging body of literature on judgment aggregation.<sup>64</sup> This literature is distinct from the extensive literature on Condorcet theorems in being mainly concerned with the *coherence* of collective judgments, rather than their *correspondence* with the facts.<sup>65</sup> As one can see from the above *Discursive paradox* case, the problem is not described in terms of the correct option (and the voters' competence or reliability to judge correctly), but solely in terms of the incoherence of the collective judgments that result from the operation of simple majority rule on coherent individual judgments.

The present study is concerned with the collective optimality of the weighted majority rule. For this purpose what is relevant is the collective's (expected) ability to judge correctly — by weighted majority rule — which option is collectively better. The above *Discursive paradox* case does not specify whether propositions  $a$  and  $b$  are in fact true or false. Let us just for the sake of the argument assume that both happen to be true. Then, as *Table 1* shows, the majority outcome will be incorrect if and only if a majority of

---

<sup>63</sup> More specifically, the set of collective judgments violates the *consistency* requirement that it must be possible for all propositions within a given set of judgments to be simultaneously true. And it violates the requirement of *deductive closure* that if the propositions within a subset of the given set of judgments logically entail another proposition, the latter must be included within the given set of judgments as well (or at least, its negation must not be included). Cf. List (2012: 207f.).

<sup>64</sup> See e.g. Pettit (2001). The discursive paradox has been generalised from what is known as the 'doctrinal paradox' concerning judges' collective decision-making in court (see Kornhauser and Sager 1986). As List (2012: 209, footnote 7) characterises these two, the 'doctrinal paradox consists in the fact that majority voting on the premises [...] may lead to a different outcome than majority voting on the conclusion', while the discursive paradox, 'more generally, consists in the fact that simultaneous majority voting on *any* set of suitably connected propositions – whether or not they can be partitioned into premises and conclusions – may yield a logically inconsistent set of judgments'.

<sup>65</sup> Cf. List (2012).

the voters judges incorrectly (as in row 3, concerning  $a&b$ ). This is rather unsurprising and not paradoxical at all. Collective incoherence does not constitute a problem when we are concerned with correspondence of collective judgments with the facts. When such incoherence is observed, we can simply infer that some of the collective judgments must be false (due to false individual judgments).<sup>66</sup>

Still, the problem should not just yet be brushed aside, as there are some interesting lessons to be drawn from *Discursive paradox* cases such as the above. Consider the following case.

**New discursive paradox.** Again, the three voters vote on two independent propositions, namely  $a$ :  $x$  is in the common interest (rather than  $y$ ), and  $b$ :  $z$  is in the common interest (rather than  $w$ ). Let us say for simplicity that each voter has an equal 0.7 probability of judging correctly, that they judge independently, and that  $a$  and  $b$  in fact both are true. Then, Classical CJT (restated in terms of the probability of correct individual and group judgments, rather than votes and outcomes) tells us that the probability of a correct majority judgment is  $P_n > 0.7$ . In fact, we can easily see that the probability of a correct majority judgment, concerning either proposition, would be  $P_n = 0.784$ .<sup>67</sup> The probability of a correct majority judgment on *both*  $a$  and  $b$ , taken separately, would hence be  $0.784^2 = 0.615$ .

However, consider now the decision on the conjunction  $a&b$ . From the individual probabilities concerning the (independent) propositions  $a$  and  $b$  we can infer that the voters have a  $0.7^2 = 0.49$  probability of correctly judging their conjunction. Then, the probability of a correct majority judgment is only  $P_n = 0.485$ .<sup>68</sup> This is significantly lower than the probability of a correct majority judgment on  $a$  and  $b$  taken separately (0.615).

It may at first sight seem paradoxical that group competence (the probability of a correct majority vote) differs for the same pair of propositions, depending on whether they are judged separately or in conjunction. Yet this result should not come as a surprise — calling it a paradox is really something of an overstatement. Consider that on each of the separate propositions, individual competence is well above chance. Yet on the conjunctive proposition, it is below chance. Hence, Classical CJT does not apply. Rather, what does apply here, and explains the poor result, is its ‘sad flipside’ of the Condorcet theorems (as referred to in section 5.2.1): that for worse-than-chance proba-

---

<sup>66</sup> Bovens (2006) shows that Wolff’s ‘mixed motivation problem’ (see above, 4.2.1 and 5.2.5) is structurally similar to, and hence can be stated as, a discursive paradox. As argued above, the alleged problem of mixed motivation is really a problem of voter incompetence. It is therefore not very surprising that the structurally similar discursive paradox, as framed in the present context with a given standard of correctness, is first and foremost problematic in terms of voter and group incompetence, rather than in terms of collective incoherence.

<sup>67</sup> There is a majority for the correct option if and only if all three, or exactly two, of the voters vote correctly ( $P_n = 0.7^3 + 3(0.7^2 \cdot 0.3) = 0.343 + 0.441 = 0.784$ ).

<sup>68</sup> There is a majority for the correct option if and only if all three, or exactly two, of the voters vote correctly ( $P_n = 0.49^3 + 3(0.49^2 \cdot 0.51) = 0.485$ ).

bilities, the non-asymptotic conclusion is that the probability of a correct majority judgment is *below* individual (or average) probability, and the asymptotic conclusion is that for increasing numbers, the probability of a correct majority judgment decreases towards *zero*.

What the *New discursive paradox* case illustrates, however, is that the nature of the decision at hand may affect individual probabilities to judge correctly, such that, even if all voters are quite competent on separate proposition, conjoining them into compound (e.g. conjunctive) propositions may result in much lower individual competences and hence much worse group results.<sup>69</sup> In such cases, it would be advisable to *decompose* compound propositions into separate ones on which voters are more competent.<sup>70</sup> In the *New discursive paradox*, decomposing *a&b* into *a* and *b* would improve group competence (from 0.485 to 0.615).<sup>71</sup>

---

<sup>69</sup> Cf. Estlund's (2008: 229) 'disjunction problem'.

<sup>70</sup> List (2012: 220–223) proposes such 'decomposition' of compound propositions as one solution for increasing correspondence of collective (majority) judgments with the facts. This is also sometimes referred to as a 'premise-based procedure' — in contrast to a 'conclusion-based procedure', see e.g. Bovens and Rabinowicz (2006). A qualifying note: Bovens and Rabinowicz (2006) show the premise-based procedure to be superior when it comes to tracking correctness 'for the right reasons' — yet, in certain circumstances, the conclusion-based procedure is superior when it comes to simply tracking correctness, for any reason. I do not go into these subtleties here, as I do not pursue the solution of a premise-based procedure (decomposition) much further.

As another solution for increasing correspondence of collective (majority) judgments with the facts, List (2012: 223–226) proposes 'decentralization': 'Instead of requiring every group member to make a judgment on every premise, the group may divide itself into several subgroups, for simplicity of roughly equal size: one for each premise, where the members of each subgroup specialize on their assigned premise and make a judgment on it alone'. Then, there will be 'epistemic gains from specialization' and 'epistemic losses from lower numbers' within the subgroups, between which there will be some trade-off (2012: 224, italics omitted). This complex voting rule, however, is far from the weighted majority rule, as analysed in the present study, and is therefore not considered any further.

<sup>71</sup> However, decomposition of compound propositions is not always advisable. Consider the following case. There are two propositions, *a* and *b*, both of which are both true. Voters  $i_1$ ,  $i_2$  and  $i_3$  each have a 0.4 reliability of judging correctly. Then, group reliability  $P_n = 0.4^3 + 3(0.4^2 \cdot 0.6) = 0.352$  for each of these two decisions. However, now we instead have all three votes on the disjunction of the two propositions, that is, *avb*. Since, as assumed, both *a* and *b* are true, *avb* is true. Each voter will judge this compound disjunctive proposition correctly if and only if she either judges both propositions correctly, or exactly one of them. Hence, individual voter reliability on *avb* is  $0.4^2 + 2(0.4 \cdot 0.6) = 0.64$ . But then, on the compound proposition of *avb*, group reliability  $P_n = 0.64^3 + 3(0.64^2 \cdot 0.36) = 0.7045$ . So in such a case, where voter reliability is higher on the compound proposition than on its separate parts, decomposition would not be advisable at all. However, for the present study, compound disjunctive propositions are of no great relevance. Spelling out propositions *a* and *b*, their disjunction states 'either *x* is in the common interest (rather than *y*) or *z* is in the common interest (rather than *w*)'. A collective judgment that this disjunction is true is not very helpful for determining whether to implement options *x* and *z* or not — which is the question we are ultimately interested in. (In contrast, a collective judgment that the compound conjunctive

Now, note that in this study I have from the outset made the simplifying assumption that the options in the collective decision — or in the present case: the propositions on which the collective judgment is made — are given. By this assumption, we have escaped complications from agenda setting, such as that those in charge of the agenda may thereby have excess influence over the outcome, apart from the influence conveyed to them through their weighted votes. (This may endanger the optimality or correctness of the outcome: consider an agenda-setter who is very keen on avoiding the collective judgment that  $b$  is true in the above *New discursive paradox* case. By having the group decide on  $a\&b$  — rather than  $a$  and  $b$  separately — she increases the odds for achieving her end, since the most likely outcome is that  $a\&b$  are collectively judged as false.) The proposal to decompose compound options, or propositions, violates the assumption that the options on the agenda are given. I therefore do not consider the proposal any further. But it is worth noting that agenda problems are relevant problems that eventually have to be dealt with.

A further issue that may arise from decomposing compound options is the problem of strategic voting. In a binary choice, a voter cannot achieve her ends (whatever these are) better by voting against, rather than for, the option that best fulfils her ends. In a non-binary choice context, where there are more than two options (as might result from decomposing the compound options of a binary choice), a voter's ends may at times be served better if she votes for an option that, in itself, does not fulfil her end. I look into the problem of strategic voting in the next chapter.

One final note: framing the compound propositions in terms of  $a\&b$ , and having voters vote for it being simply true or false, may also be faulted for obfuscating some possibly relevant details. Consider that for  $a\&b$  to be true, both  $a$  and  $b$  must be true. But for  $a\&b$  to be false, there are three separate possibilities: either  $a$  is false, or  $b$  is false or both  $a$  and  $b$  are false. So there are really four different possibilities underlying the binary judgment that  $a\&b$  is true or false.<sup>72</sup> We could now reframe the decision in sufficient detail and have voters vote for which one of the following mutually exclusive and jointly exhaustive propositions is true:  $a\&b$ , not- $a\&b$ ,  $a\&\text{not-}b$ , and not- $a\&\text{not-}b$ . Then, a voter with a 0.7 probability of judging that  $a$  and  $b$  (separately) are true will have a seemingly poor 0.49 probability of judging that  $a\&b$  is true. Yet this figure is much higher than any one of the voter's probabilities on the remaining three propositions: 0.21 on not- $a\&b$  and  $a\&\text{not-}b$ ,

---

proposition 'x is in the common interest (rather than y) and z is in the common interest (rather than w)' is true can reasonably be used for this purpose, if the group seeks to implement common-interest options.)

<sup>72</sup> Spelling out the propositions  $a$  and  $b$  in terms of what is in the common interest,  $a\&b$  is true when  $x$  and  $z$  are in the common interest (rather than  $y$  and  $w$ ). Yet  $a\&b$  is false when  $y$  and  $z$  are in the common interest (rather than  $x$  and  $w$ ), or when  $x$  and  $w$  are in the common interest (rather than  $y$  and  $z$ ) or when  $y$  and  $w$  are in the common interest (rather than  $x$  and  $z$ ).

respectively, and 0.09 on not-*a*&not-*b*. If we were to apply plurality rule (where voters cast their vote for exactly one out of  $m > 2$  options, and the option receiving most votes is selected as collective outcome), *a*&*b* would be most likely to be selected.<sup>73</sup>

It would be interesting to see how plurality rule could be redesigned to operate on weighted votes, whether the rule then could be shown to be (weakly) collectively optimal and whether it would outperform the weighted majority rule. I do not pursue these questions further within this study, which is limited to examining the weighted *majority* rule. Still, they are well worth taking up on another occasion.

## 6.5 Conclusions

In this chapter, one of the previous chapter's arguments from weak collective optimality has been restated conditional on an assumption of voter behaviour — regardless of how this may be motivated. The *behavioural argument from weak collective optimality* showed that the weighted majority rule is weakly collectively optimal if votes are on average minimally reliable to be cast for the correct option, given that they are sufficient in (odd) numbers, and on average tolerably correlated.

This result can now be related to the extended argument from collective optimality, which was stated conditional on Self- Or Common-Interested Voting. The now stated Minimal Average Reliability assumption can be understood to relax Self- Or Common-Interested Voting, since it allows any kind of voting behaviour (self-interested, common-interested, partial-interested, etc.) — constrained by the requirement that the votes are cast for the common-interest option with a better-than-chance probability. Relaxing Self- Or Common-Interested Voting has had the price of introducing an additional assumption (of sufficiently low average correlation) and of a result in terms of weak (rather than strong) collective optimality.

The arguments of Chapter 5 have shown under which conditions the weighted majority rule is weakly collectively optimal. One crucial condition has been that there be a sufficiently large (odd-numbered) group of voters or votes. This means that within the traditional political domain of elections and referenda, with large electorates, the weighted majority rule emerges as a promising candidate from a common-good perspective. Yet for small-scale collective decision-making, weak collective optimality could be preserved only for much higher CJT-competence levels. However, I have proposed an alternative argument for the weighted majority rule. The *better-than-the-average-voter argument* established that, for on average minimally CJT-

---

<sup>73</sup> Cf. List and Goodin (2001: 279) who argue 'that the Condorcet jury theorem can indeed be generalized from majority voting over two options to plurality voting over many options'.

competent voters, the weighted majority rule is more likely than the average voter to select the common interest option, if correlation between votes is constrained appropriately. However, I have conceded that performing better than the average voter is not a very recommending property of the weighted majority rule.

A final brief consideration of the discursive paradox gave at hand that the way the options are set up on the agenda — e.g. as compound or separate option — may have repercussions on voter competence, and hence, on the weak collective optimality of the weighted majority rule. This section introduced non-binary choice contexts, where there are more than two options on the agenda. Such multi-option contexts are the field of inquiry for my next chapter.



## 7 Multi-option decisions

### 7.1 Introduction

In the previous chapters, I have focused on binary decisions, that is, decisions with exactly two options. Now I want to expand the scope of the investigation and consider the weighted majority rule's performance in multi-option decisions, that is, in collective decisions with more than two options. I specifically examine whether the case for the weighted majority rule is undermined by the possibility of strategic voting that may arise for self- or common-interested voters. I say that a voter *votes strategically* in a particular decision if and only if she votes against her desired end (be it her self- or the common interest) in this decision, to promote this end on the whole. Strategic voting is not possible in binary decision contexts, that is, when 'the whole' consists of only one decision with two options. Then, promoting one's end on the whole means voting according to this end in the particular decision. In multi-option decisions, this might be a different matter.

Strategic voting poses a threat to the case for the weighted majority rule if (1) strategic voting is possible under this rule and (2) it leads to collectively suboptimal outcomes. It may now be objected that there is not much to investigate then since it has long been established, regarding the first claim, that there are *no* strategy-proof voting rules that satisfy some minimally reasonable conditions, and regarding the second claim, that strategic voting has damaging effects on the efficiency of voting rules. Let us look at the objection in some more detail.

Regarding claim (1), one may refer to the Gibbard-Satterthwaite theorem, which establishes that, for decisions with more than two options, *no* non-dictatorial, predictable (that is non-random) voting rules are strategy-proof. A voting rule is *strategy-proof* if under it no voter has an incentive to vote strategically (as defined above). More specifically, Allan Gibbard shows 'that any non-dictatorial voting scheme with at least three possible outcomes is subject to individual manipulation'.<sup>1</sup> However, we should note that Gibbard specifies a 'voting scheme' as 'any scheme which makes a community's choice *entirely* dependent on individuals' professed preferences among the

---

<sup>1</sup> Gibbard (1973: 587).

alternatives'.<sup>2</sup> As such, voting schemes do not allow any involvement of chance in the determination of the collective outcome, as from a random tie-breaker.<sup>3</sup> In fact, Gibbard requires any voting scheme to have one unique collective outcome for any profile of individual preferences, so top ties between options are ruled out as well. Likewise, Mark Satterthwaite, in his proof that a strategy-proof voting rule must be dictatorial, considers exclusively a 'committee [which by means of some voting rule] selects a *single* alternative from the alternative set'.<sup>4</sup> Thus, he does not allow top ties either.

As we have seen in binary decision settings, the weighted majority rule may result in ties whenever both options are collectively equally good. In such cases, we have seen, the case for the weighted majority rule is compatible with any kind of tie-breaker, as either option is as good as the other. As we will see shortly, the rule does allow top ties in multi-option settings as well. Thus, the Gibbard-Satterthwaite theorem does not apply to our present context.

Allan M. Feldman extends Gibbard's and Satterthwaite's results to voting rules that allow top ties and resolve them by an even-chance lottery. More specifically, he shows that any such voting rule 'must be a dictatorship [rule of one voter] or a duumvirate [equally probable rule of one of two voters]' if it is strategy-proof and 'non-imposed' (where non-imposition simply means that, for any subset of all the given options, there is a profile of individual rankings under which this subset is at the top of the collective ranking).<sup>5</sup> Crucially, however, Feldman's result hinges on the assumption of an even-chance lottery as a tie-breaker, whose assignment of probabilities to the tying options within the top-ranked subset is common knowledge.<sup>6</sup> This raises

---

<sup>2</sup> Gibbard (1973: 587, italics added).

<sup>3</sup> Gibbard calls any system that allows random tie breakers a 'mixed decision scheme' rather than a voting scheme. One such 'mixed decision scheme' is a voting lottery (where e.g. a random vote is drawn from an urn to determine the outcome) that is easily shown to be strategy-proof (cf. Gibbard 1973: 593; cf. also Satterthwaite 1975: 193, footnote 6).

<sup>4</sup> Satterthwaite (1975: 190, italics added).

<sup>5</sup> Feldman (1980: 509).

<sup>6</sup> Duggan and Schwartz (2000) generalise the Gibbard-Satterthwaite and Feldman theorems to apply to voting rules that allows top ties to be resolved by *any* random tie-breaker, even if the involved probabilities are not common knowledge. But their theorem rests on a condition that is not satisfied under weighted majority rule: 'If all [individual rankings]  $P_{j \neq i}$  are the same, with  $x$  first and  $y$  second, and if  $P_i$  is either the same as them or else the same but with  $y$  first and  $x$  second, then [the collective winner] is a singleton [that is, one unique option]' (2000: 87). This condition thus prevents the theorem to apply to any voting rule which, for such a distribution of individual preference orderings, would allow a top tie (to be resolved by some tie-breaker). The weighted majority rule may, however, generate a top tie in such a case, as can be seen from the following. If everyone but  $i_j$  rank  $x > y$ , while  $i_j$  ranks  $y > x$ , and if  $i_j$ 's stakes are as great as the sum of everyone else's stakes in this decision, there is a tie between  $x$  and  $y$ . And if everyone (including  $i_j$ ) ranks all other options below  $x$  and  $y$ , the tie between  $x$  and  $y$  is at the top of the collective ranking. This case violates Duggan and Schwarz's condition, and hence their theorem does not apply to the weighted majority rule.

the question whether there might be other tie-breakers that could rule out strategic voting. Thus, there is room for some further investigation, concerning assumptions about tie-breakers and other conditions that might ensure strategy-proofness.<sup>7</sup> This is one line of inquiry for the present chapter.

A separate line of questions concerns claim (2) above, namely, whether and to what extent strategic voting, when possible, undermines the collective optimality of the weighted majority rule. When it comes to simple majority rule, some writers indeed emphasise the collectively damaging effects of strategic voting.<sup>8</sup> Others, however, view it as collectively beneficial. Buchanan and Tullock make explicit the optimising effects of *logrolling*.<sup>9</sup> Logrolling is strategic voting practiced by coalitions of voters, over several decisions, in which coalition members ‘trade’ votes with each other. Buchanan and Tullock argue that permitting logrolling under simple majority rule can, for instance in a decision that strongly affects a minority and insignificantly affects an opposed majority, ‘result in a great increase in the well-being of both groups, and the prohibition of such transactions will serve to prevent movement toward the conceptual “social optimality” surface, under almost any definition of this term’. They argue that thus ‘a reasonably strong ethical case can be made for a certain amount of vote-trading under majority-rule institutions’.<sup>10</sup>

However, with the weighted majority rule, within my proposed framework, strategic voting cannot be collectively beneficial — that is, collectively better than non-strategic voting. Consider that the further-extended argument from collective optimality has shown the weighted majority rule to select collectively optimal outcomes, when voters vote according to their self- or the common interest (that is, non-strategically). Now, a voter can be assumed to vote strategically only if this will result in an outcome which is better — in terms of her desired end — than the outcome from non-strategic voting. Yet such an outcome cannot be better — in terms of collective optimality — than the collectively best outcome, which is what results from non-strategic voting. So strategic voting cannot be collectively beneficial. Still, it is an open question whether and to what extent strategic voting under the weighted majority rule is collectively damaging. Can it result in the selection of collectively worst options — or can it only achieve a slightly worse than collectively optimal outcome? Or, even less severe, can strategic voting only

---

<sup>7</sup> Note that under the assumption that individual preferences are ‘single-peaked’ on one dimension along which all options can be ordered (e.g. the usual political left–right dimension), it can be shown that there are non-dictatorial, predictable voting rules that are strategy-proof. (Cf. Barberà 2011: 759–760; Black 1958.) I return to this possibility in 7.6 below.

<sup>8</sup> See e.g. Riker (1982: 161–165). For a good discussion of Riker’s claims, see Mackie (2003: 163–166).

<sup>9</sup> Buchanan and Tullock (2004: chapters 9–10).

<sup>10</sup> Buchanan and Tullock (2004: 129).

ensure specific outcomes among the options that tie for the top of the (non-strategic) collective ranking?

Strategic voting is frequently demonised as ‘manipulating the outcome’, as ‘strategically misrepresent[ing] preferences’, as ‘sophisticated’ voting, and the like.<sup>11</sup> The allegation seems to be that there is some genuine individual ranking of the options which is misrepresented or insincerely reported for some dubious purpose. But this is out of place, as Michael Dummett points out.

‘A voting paper is not [...] a questionnaire. It is a mechanism in a decision-making procedure which will have whatever effect it has because of the way in which that procedure works. The actual significance of a particular vote is therefore wholly determined by the procedure of which it forms part [...]. The only question that a voter actually answers by casting a particular vote is, ‘In which way do you think that you should cast your vote in order to obtain an outcome as agreeable to you as possible?’, and this question he inevitably answers sincerely’.<sup>12</sup>

This chapter will give us a better understanding of Dummett's point.

I propose a straightforward way of employing the weighted majority rule to multi-option decisions in 7.2 and show that the rule is collectively optimal even in this extended context, under the Self- Or Common-Interested Voting assumption. I then, however, point out that this assumption is ambiguous in multi-option contexts. Depending on how the assumption is interpreted, the case for the weighted majority rule may be undermined due to strategic voting. I start out by considering self-interested and fully competent voters. In 7.3 I explore some of the possibilities — and limits — of strategic voting, when this is practiced by *groups* of self-interested voters (‘logrolling’). Section 7.4 explores some of the possibilities — and limits — of strategic voting, when this is practiced by *individual* self-interested voters (‘individual strategies’). The question of whether the weighted majority rule could be rendered in alternative versions that avoid these problems is briefly discussed in 7.5. In 7.6 I consider strategic voting with common-interested voters. Section 7.7 considers less than fully competent voters. Finally, section 7.8 concludes.

## 7.2 The further-extended argument from collective optimality

Consider the following case, which closely resembles *Dinner plans 3* (in 2.4.2 above), but with a slightly different distribution of stakes. (For the sake

---

<sup>11</sup> Mackie (2003: 35, 39, 41).

<sup>12</sup> Dummett (1984: 79–80).

of simplicity, all examples in this chapter are formulated for a sum-total criterion and stakes are, straightforwardly, individual well-being differentials. The examples could, however, be modified to encompass such functions.)

**Dinner plans 4.** There are three people,  $i_1$ ,  $i_2$  and  $i_3$ . They are to decide whether to go *out* to a restaurant, stay *in* and cook or just *order* dinner on the phone. Each of them is affected differently by the three options. The individual rankings of the three options, as well as the individual stakes for each pair of options, can be read off from *Table 1*. Again, the exact figures in each cell are not important. What matters is that, the higher they are, the better the option for the individual. Moreover, the differences between the figures give us the (interpersonally comparable) stakes for each individual. Given these stakes, we can read off that the third option *order* is the collectively best option.

	$i_1$	$i_2$	$i_3$
<i>in</i>	2	0	1
<i>out</i>	1	2	0
<i>order</i>	0	1	3

*Table 1*

Could these voters use the weighted majority rule? Would it select the common-interest option, *order*? In the previous chapters the rule has been defined for binary decisions (see 2.2 above).

**The Weighted Majority Rule:** For all individuals and any decision with two options, (a) every individual is assigned a number of votes in proportion to her stakes, and (b) the option that receives a majority of votes is selected as outcome.

From this, it is not immediately obvious how the weighted majority rule would operate in non-binary decisions.

Let us briefly go back to Brighthouse and Fleurbaey's argument for the weighted majority rule in the light of a prioritarian criterion of collective optimality (see 3.2.1 above). We are invited to suppose 'that the options are ranked by application of the weighted majority rule *over every pair of options*'.<sup>13</sup> This indicates that their argument is not limited to singular binary decisions, but could also be applied to decisions with more than two options. The quoted phrase could be taken to instruct us on how to use the weighted majority rule for decisions with more than two options: each of the options is run pairwise against every other option, and for each pair, the weighted majority rule is applied as before. From the resulting pairwise rankings of the options, a collective ranking is then derived with a simple *ordinal ranking*

<sup>13</sup> Brighthouse and Fleurbaey (2010: 144, my italics).

*rule*: an option  $x$  is collectively ranked above an option  $y$  if and only if  $x$  receives more votes than  $y$  in a pairwise vote. Finally, the top-ranked option is selected as the outcome or ‘winner’ of the multi-option decision. This procedure is reminiscent of the so-called Condorcet rule, which in the same way extends *simple* majority rule to multi-option decisions (cf. 2.4.1 above). Let us therefore call this a *Condorcet-style* extension of the weighted majority rule and define it as follows.

**The Condorcet-Style Weighted Majority Rule:** For all individuals and any decision, for each pair of options, (a) every individual is assigned a number of votes in proportion to her stakes for this pair, (b) an option  $x$  is collectively ranked above an option  $y$  if and only if  $x$  receives more votes than  $y$  in a pairwise vote, and (c) the top-ranked option is selected as the outcome.

(Note that in this chapter, when referring to the weighted majority rule I refer to its now defined Condorcet-style extension.) We can now apply this rule to the above case.

**Dinner plans 4 (continued).** There are three pairs of options (*in-out*; *in-order*; *out-order*) on which the three voters can vote. When applying the Condorcet-style weighted majority rule, votes are distributed in proportion to the individual stakes for each pair of options. So, e.g. in the decision *in-out*,  $i_1$  and  $i_3$  get one vote each, while  $i_2$  gets two votes. Let us assume that all vote according to their self-interest (as reflected by their stakes). Then,  $i_1$  and  $i_3$ , with one vote each, vote against  $i_2$ , with two votes, so there is a tie. Let us express this as ‘*in-out*’. On *in-order*,  $i_1$  and  $i_3$  get two votes each, and  $i_2$  one vote. The former two vote against each other, and so the latter decides with her vote: *order* is ranked above *in*. Let us express this as ‘*order > in*’. Finally, on *out-order*,  $i_1$  and  $i_2$  get one vote each, while  $i_3$  gets three and therefore decides in favour of *order*, such that *order > out*. Then, we can construct a collective ordinal ranking, such that *order > in-out*. The top-ranked option is *order*; it is therefore selected. And this means that the collectively best option is selected.

The example suggests that the collective optimality of the weighted majority rule extends to its Condorcet-style extension.

When it comes to the *simple* majority rule, the worry with such a Condorcet-style extension is that it might produce a cyclic collective ranking of the options. Given three options  $x$ ,  $y$  and  $z$ , it might turn out that  $x$  beats  $y$ ,  $y$  beats  $z$  and  $z$  beats  $x$ , by a ‘rotating’ majority of votes (as illustrated with the example of the French presidential election, in 1.1 above). Now it is interesting to note that, in the typical context of *Classical CJT*, with correctness-tracking voters, cycling must be due to voter incompetence (or unreliability).<sup>14</sup> Were all voters fully competent (reliable), all would vote according to

---

<sup>14</sup> Cf. 5.2.1 above. For a good discussion and re-interpretation of Condorcet's own proposal on how to resolve such cycles (namely by disregarding the pairwise ranking(s) that have the

the same, correct — presumably non-cyclical — ranking of the options, and cycling would be impossible. However, once we allow voters to have heterogeneous ends — to be self- or common-interested — cycles may result even when all voters are perfectly competent (reliable) with regard to these ends. Consider again *Dinner plans 1* above, and assume that the three voters are self-interested and perfectly competent (that is, vote reliably according to their self-interest). If they receive *one* vote each on every pairwise decision, *in* would beat *out*, with  $i_1$ 's and  $i_3$ 's two votes, *out* would beat *order*, with  $i_1$ 's and  $i_2$ 's two votes — and *order* would beat *in*, with  $i_2$ 's and  $i_3$ 's two votes. So, with simple majority rule, cycling is possible even with perfect competence (reliability), when voters do not share the same goal.

Yet we can easily see, from the *generic argument from collective optimality* (see 3.3 above), that cycling under the *weighted* majority rule is impossible in such cases — as long as voters vote according to their self-interest. If this is the case, the argument showed that, in a binary decision, the weighted majority rule ranks the collectively better option higher than the other option (and makes equally good options tie). Since this holds for all pairwise (that is, binary) decisions, into which a multi-option decision may be decomposed, the rule perfectly tracks how all the options are ranked according to the criterion of collective optimality. Since such a ranking cannot reasonably be cyclic, the options themselves cannot cycle either.<sup>15</sup>

This also means that even in multi-option decisions, the Condorcet-style weighted majority rule must rank the collectively best option at the top of the collective ranking — if all voters vote according to their self-interest, that is, according to Self-Interested Voting. Now, imagine that one of the voters, rather than voting for her self-interest option, votes for the common-interest option in a particular pairwise decision. Then, the group of voters satisfies the relaxed Self- Or Common-Interested Voting assumption. Now, such a change in assumed voting behaviour could never *decrease* voter support for the collectively better option within the pair, which means that it cannot reverse the ranking of the two options. The weighted majority rule must then still select the same, collectively best option. This argument can be made for any number of voters that switch from self- to common-interest voting, in any number of the pairwise decisions at hand. Thus, the *extended argument from collective optimality* (see 4.2 above) can be easily further extended, from binary to multi-option decisions.

I call the just stated argument the *further-extended argument from collective optimality*. Its conclusion states that the weighted majority rule is collec-

---

weakest majority support), see Young (1988). I briefly return to cycling due to voter incompetence in 7.5 below. Cycling may also be possible due to strategic voting behaviour, see 7.3–7.4 below.

<sup>15</sup> Cf. Brighthouse and Fleurbaey (2010: 144); Fleurbaey (*mimeo*: 30).

tively optimal even in the latter context, given Self- Or Common-Interested Voting.

However, we need to tread carefully at this point. In multi-option settings, it is not quite clear exactly how Self- Or Common-Interested Voting should be understood. I bring out this difficulty in the next section.

### 7.3 Logrolling for self-interested voters

Consider the following two cases, where four options are on the table.

**Dinner and a movie 1.** A group of three,  $i_1$ ,  $i_2$  and  $i_3$ , needs to make two decisions. First, these three need to decide whether to eat *in* or *out* tonight. Second, they need to decide whether to watch a *comedy* or a *horror* film afterwards. *Tables 2* and *3* spell out the voters' stakes in the respective decision. Let us assume that all vote according to their self-interest in each of the two decisions.

As it happens, voter  $i_3$  is pivotal in the first decision with options *in* and *out*. (As stated before, a voter is pivotal in a binary decision if and only if, for a given distribution of all the voters' votes between the two options, had this voter voted for the other option, the outcome would have changed to that option.)  $i_3$  ranks *out* > *in*, but her stakes are quite low (though not zero). There is some other voter,  $i_2$ , who ranks *in* > *out*, and whose stakes are quite high in this decision — say that she is very hungry but broke. Unfortunately for her,  $i_2$  is not pivotal in this decision. In the second decision between *comedy* and *horror*, the tables are turned. That is,  $i_2$  who ranks *comedy* > *horror* is pivotal but has low stakes. While non-pivotal  $i_3$  has quite high stakes — she regularly suffers from severe anxiety attacks after watching comedy movies and thus ranks *horror* > *comedy*. Since everyone votes according to their self-interest (as straightforwardly reflected by their stakes) in each decision, the collectively optimal options *out* and *comedy* are selected by the weighted majority rule (due to pivotal  $i_3$  and  $i_2$ , respectively). Then, both  $i_2$  and  $i_3$  get what they want in their respective low-stake decision — but fail to get what really matters to them.

	$i_1$	$i_2$	$i_3$
<i>out</i>	5	0	1
<i>in</i>	0	5	0

*Table 2*

	$i_1$	$i_2$	$i_3$
<i>comedy</i>	5	1	0
<i>horror</i>	0	0	5

*Table 3*

However, one might wonder, would it not in fact be better for  $i_2$  and  $i_3$  to vote in another way? Consider the following.

**Dinner and a movie 2.** We are looking at the same group of voters, with the same stakes in the same pair of binary decisions. But now,  $i_2$  and  $i_3$  'help each other out': they engage in *logrolling*. In the first decision, pivotal voter  $i_3$  casts her one vote with  $i_2$  for *in*, and in the second, pivotal voter  $i_2$  casts her one vote with  $i_3$  for *horror*. By doing so, they sacrifice a little in their respective

low-stake decision (getting one unit less than they would have otherwise), yet gain more in their respective high-stakes decision (getting five units more). Then, all in all, they are made better off. Thus, we can say that  $i_2$  and  $i_3$  vote according to their self-interest, on the whole, by logrolling. In this case, however, the weighted majority rule selects the collectively suboptimal options *in* and *horror*.

In both scenarios it is stated that  $i_2$  and  $i_3$  vote according to their self-interest, such that they satisfy Self- Or Common-Interested Voting. Yet this implies different forms of voting behaviour in the two cases. And it ensures collective optimality of the weighted majority rule in the first — but not in the second — scenario. This suggests that Self- Or Common-Interested Voting is ambiguous in multi-option settings. And to be sure, *Dinner and a movie 1* states that the voters vote according to their self-interest *in each of the two decisions*, while *Dinner and a movie 2* states that they vote according to their self-interest *on the whole*. So the assumption has different scopes: let us say that it has *local scope* when it applies to each binary decision in isolation, and *global scope* when it applies to an entire multi-option decision (consisting of the set of all binary decisions on which votes may be traded). Now we can disambiguate the following two versions of Self- Or Common-Interested Voting (or SCI-Voting for short).

**Local-Scope SCI-Voting:** Every voter votes for the option that is in her self-interest or in the common interest, *within each binary decision*. That is, every voter votes according to her *local* self-interest or to the *local* common interest.

**Global-Scope SCI-Voting:** Every voter votes such as to promote the option that is in her self-interest or in the common interest, *among all the available options*. That is, every voter votes according to her *global* self-interest or to the *global* common interest.

The above *further-extended argument from collective optimality* was conducted under the Local-Scope interpretation of SCI-Voting. To be precise, the argument relied on the claim that the weighted majority rule ranks the collectively better option above its alternative in each pairwise (binary) decision. And this claim, in turn, had been shown (by the *extended argument from collective optimality*) to hold for each binary decision, if voters vote for their self-interest or the common-interest option *in this decision* (see 4.2 above). This condition is equivalent to Local-Scope SCI-Voting.

So when arguing for the weighted majority rule's collective optimality in multi-option decisions, one should be clear that what is assumed is Local-Scope SCI-Voting. This point can also be directed at Brighthouse and Fleubaey's argument for the weighted majority rule. Recall that their argument can be read as applying to multi-option decisions, where 'the options are ranked by application of the weighted majority rule over every pair of op-

tions'.<sup>1</sup> The argument rests on the assumption 'that every individual always votes according to his interests'. This could easily be taken to be equivalent to my above Global-Scope SCI-Voting — which, as we have just seen, cannot establish the conclusion Brighthouse and Fleurbaey aim for. Thus, they need to precisify this crucial premise.

Fleurbaey's separate argument for the weighted majority rule instead rests on the assumption 'that each voter votes for the option which is more favorable to him'.<sup>2</sup> This quite closely resembles my above Local-scope SCI-Voting (when the latter is limited to self-interested voting and Fleurbaey's term 'favourable' is spelled out in terms of self-interest).<sup>3</sup> However, even Fleurbaey could be clearer about the implication of this assumption, namely that *voting for the option* that is more favourable to a voter may at times be less favourable to her than voting against it.

The problem with this implication is that it renders the case for the weighted majority rule somewhat less appealing than what was hoped for from the outset. Recall that the proclaimed merit of this rule (as stated in 3.3 above) is that it manages to derive the winning common-interest option, as an output, from input information in terms of what is in the voters' self-interest. However, it now turns out that in multi-option settings, 'what is in the voters' self-interest' must be understood in a rather myopic sense, and may occasionally be against their self-interest on the whole. Since (as claimed in 2.3.1 above) any plausible theory of individual well-being is concerned with how people fare on the whole — over a whole life or at least a considerable stretch of time, rather than in some more or less arbitrary myopic sense — this means that what is really assumed here is that people occasionally vote *against* their self-interest. So it would be less misleading to say that, in multi-option settings, the weighted majority rule derives the winning common-interest option from input information in terms of what is (occasionally) against the voters' self-interest.

This is of course not an unusual position within normative political philosophy. The individual is required to sacrifice, or at least to disregard her self-interest, for the benefit of the collective. But in the present context this position implies that the weighted majority rule derives the winning common-interest option from elicited information on *either* what is in voters' self-interest *or* what is against voters' self-interest — depending on whichever

---

<sup>1</sup> Brighthouse and Fleurbaey (2010: 143f.)

<sup>2</sup> Fleurbaey (*mimeo*: 6).

<sup>3</sup> Recall that when I formulated Self-Interested Voting in 3.2 above, I took Brighthouse and Fleurbaey's (2010) and Fleurbaey's (*mimeo*) assumptions to be equivalent. In the context of binary decisions, this equivalence postulate is unproblematic. Yet when there are more than two options on which voters make pairwise decisions, the two assumptions may acquire different meanings, and may contradict each other.

er makes the rule perform better in the circumstances. This risks rendering the case for the weighted majority rule rather trivial.<sup>4</sup>

Previously, I stated that Self- Or Common-Interested Voting is implied by the conjunction of the following three claims (see 4.3.1 above).

***Self- Or Common-Interested Motivation:*** Every voter is either self-interested, in the sense that she desires to promote her self-interest, or common-interested, in the sense that she desires to promote the common interest.

***Competence:*** Every voter has a correct belief that voting for some option  $x$ , among the given alternatives, promotes her self-interest (if she is self-interested) or the common-interest (if she is common-interested).

***Success:*** Every voter acts according to the belief-desire pair referred to in Self- Or Common-Interested Motivation and Competence, respectively.

In binary settings Competence is implied by:

***Competence 1:*** Every voter correctly ranks the *options* in the light of her self- or the common interest, and

***Competence 2:*** Every voter believes that voting for the higher-ranked *option* promotes her self- or the common interest.

As the above *Dinner and a movie* cases suggested, in multi-option settings a voter such as  $i_2$  who satisfies Competence 1 and Competence 2 may have an incorrect belief about which way of voting promotes her self- (or the common) interest. In these settings, Competence is rather implied by:

***Competence 3:*** Every voter correctly ranks her voting *strategies* by ranking their outcomes in the light of her self- or the common interest, and

***Competence 4:*** Every voter believes that voting according to the higher-ranked *strategy* promotes her self- or the common interest.

In the rest of this chapter, unless stated otherwise, I take Self- Or Common-Interested Motivation, Competence (as implied by Competence 3 and Competence 4) and Success for granted and examine the consequences for the collective optimality of the weighted majority rule. I use the terminology of voting ‘strategically’ or ‘non-strategically’. Voting strategically simply means voting against one’s self-interest in some isolated binary decision, in

---

<sup>4</sup> To be clear: the problem is not that it is implausible that people would vote in such a way. All around us there are many instances of people doing what is good for them in some myopic sense, even when this is worse for them on the whole. They skip their dentist appointment even though this will result in a severe toothache in the long run. They light the next cigarette in the face of the commonly known health risks. And so on. So it might analogously be true that people (occasionally) vote short-sighted and eventually against their self-interest on the whole. The problem with this assumption is instead that settling for it will leave the case for the weighted majority rule much less appealing.

order to vote according to one's self-interest the whole, thus violating Local-scope SCI-Voting. Voting non-strategically means voting according to one's self-interest in such an isolated decision. Note that voting non-strategically need not violate Global-scope SCI-Voting. For instance, it may simply be better for some self-interested voter  $i$ , on the whole, to vote according to her self-interest in a binary decision. Or it may be part of some larger strategic plan. Suppose, for instance, that  $i$  engages in logrolling with some other voter, agreeing to vote strategically to help the other one out and be helped in return. But then  $i$  defects from the deal and votes 'non-strategically' after all — because defecting is better for her. This consideration brings out the question whether a logrolling coalition can be stable, that is, secure from defecting by its self-interested members.

### 7.3.1 Is logrolling eventually against the logrollers' self-interest?

It may be objected that logrolling is not really an issue — at least under the current idealised settings, where voters satisfy assumptions Self- Or Common-Interested Motivation, Competence, and Success. The objection goes that, even though logrollers *collectively* might gain by cooperating on some decisions, *each* of them will gain by defecting from such cooperation — whatever the others do. Consider again *Dinner and a movie 2*. We assumed pivotal  $i_3$  to have low stakes in the decision with options *out* and *in*, and non-pivotal  $i_2$  to have quite high stakes. In the decision with options *comedy* and *horror*, the opposite holds. Pivotal  $i_2$  is assumed to have low stakes, and non-pivotal  $i_3$  quite high stakes. They make a deal, such that  $i_3$  votes with  $i_2$  on where to eat and  $i_2$  votes with  $i_3$  on what film to watch.

However, since by assumption  $i_3$  is self-interested and competent, she has no reason to keep her end of the bargain when she is pivotal. Consider the following: if  $i_2$  does keep her end of the deal,  $i_3$  will get her top option in *both* decisions if she defects from logrolling. If  $i_2$  does not keep her end of the deal, at least by defecting  $i_3$  will secure her top option when she is pivotal. Either way, defecting makes her better off. So  $i_3$  defects. Of course, the same applies to  $i_2$ , *mutatis mutandis*. So she defects as well, and all logrolling collapses. This means that each of them ends up worse off than she could have been, had both cooperated. This *logrollers' dilemma* is a typical prisoners' dilemma situation: regardless of whether the other cooperates (logrolls) or not, either one of them will gain by defecting and thus either one of them is worse off than she could have been. The payoffs from logrolling and defecting that result from the above *Dinner and a movie* cases (*Tables 2 and 3*) are given in the prisoners' dilemma matrix in *Table 4*.

		$i_2$	
		<b>Logroll</b>	<b>Defect</b>
$i_3$	<b>Logroll</b>	5	6
		5	0
	<b>Defect</b>	0	1
		6	1

Table 4

Thus, *for the (potential) logrollers*, the outcome of their self-interested strategies of defecting is collectively suboptimal. Of course, in another sense, the outcome — from non-strategic voting — actually is collectively optimal *for the entire group of voters*, which also includes  $i_1$ . So one could argue that, since logrolling must collapse, the case for the weighted majority rule can be restored — thanks to the logrollers' dilemma. We do not even have to devise compound options from all options on which votes may be traded to mitigate the damaging effects of logrolling. With the given binary decisions with simple options, logrolling just is not stable.

But what if more decisions were coming up in a near future? It might appear that  $i_2$  would be better off, in the long run, by cooperating (logrolling) in a given decision, assuming that (she knows that)  $i_3$  would help her out and vote with her in a later decision only if she votes with  $i_3$  in the present decision. However, if the sequence of decisions is finite, and if the logrollers are self-interested, competent and successful (according to Self- Or Common-Interested Motivation, Competence, and Success) and know all this, and know that they know it, etc., then logrolling must collapse from the outset anyway, according to the following line of reasoning. Each logroller knows that the other one would defect in the last decision in the sequence, when there is no more future cooperation to be expected. They can easily see this from the above *logrollers dilemma* argument. But then each also knows that in the next-to-last decision there is no more future cooperation to be expected and would thus defect. So the same applies to the next-to next-to-last decision, and so on, all the way until we reach the first decision on which the logrollers might cooperate, but self-interestedly choose to defect. Thus, logrolling collapses from the outset.

Moreover, this argument can be generalised to indefinite sequences of decision-making:<sup>5</sup> even if the voters are not certain about exactly how many decisions lie ahead, as long as all know (and know that they all know, and so on) that there is some definite upper limit  $n$  to the number of decisions in the sequence and that their decision to cooperate or defect does not affect this upper limit, they can employ a similar line of reasoning. They know (and know that all know, and so on) that there will be no decision, and thus no cooperation, beyond the upper limit, so on the  $n^{\text{th}}$  decision, if it ever arrives, they will defect. Then, they also know that the  $n-1^{\text{st}}$  decision, if it comes

---

<sup>5</sup> See Kavka (1986: 129–132).

about, will either be the last, or the next-to-last decision, in which no more future cooperation to be expected, so they defect. And so on, until we reach the first decision. Again, logrolling collapses from the outset.<sup>6</sup>

These *backward induction arguments*, for finite or indefinite sequences of decision-making, hinge on all voters being self-interested, competent and successful, together with the following assumption.

**Common Knowledge:** Every voter knows, and knows that everyone knows, and so on, (a) about the upper limit — or end — of the sequence of decisions they face and (b) about every voter's satisfying Self- Or Common-Interested Voting — by being self-interested — as well as Competence and Success.

So given this — rather demanding — assumption (together with the three other ones), logrolling must collapse. The collective optimality of the weighted majority rule can then be restored in multi-option decisions.

However, given some alternative assumptions, this might not be so. One possibility is to introduce some level of uncertainty in part (a) of the just stated Common Knowledge assumption.<sup>7</sup> Voters are uncertain as to the upper limit or end of the sequence of decisions on which they might logroll. However, it is common knowledge, prior to each decision, that there will be yet another decision after that, with some (constant) positive probability, e.g.  $p = 0.6$ . And it is common knowledge that everyone is self-interested, competent and successful, *and* has exactly two available strategies: 'Never logroll' — also known as 'Always Defect' — or 'Logroll initially, but respond to defection by ceasing to logroll' — also known as 'Trigger'. Then, depending on the payoffs in each decision, following the second strategy might be better for each throughout, such that logrolling would be stable from the outset.<sup>8</sup>

Consider, e.g. the above logrollers' dilemma, as specified in *Table 4*. If this decision would reoccur over and over, with a probability of 0.6, the payoffs from *Always Defect* and *Trigger*, over the sequence of  $n$  decisions, would approach the values given in *Table 5*, as  $n$  approaches infinity.<sup>9</sup>

---

<sup>6</sup> For a critical discussion of Kavka's assumption concerning common knowledge of a definite upper limit, see Jiborn and Rabinowicz (2003: 138–147).

<sup>7</sup> The outline of this argument goes back to Skyrms (2001: 33f.).

<sup>8</sup> For a possible argument to the contrary, see Jiborn and Rabinowicz (2003: 147–151). They provide the 'base step' from which to conduct a backward induction argument, by changing part (a) of Common Knowledge in the following way. It is common knowledge that the probability of there being yet another decision to come is not constant but diminishing over time. Then, at some point, it is sufficiently small to make the expect gains from cooperation in the next decision too small. Logrolling collapses at this point. Then, by backward induction, this collapse will appear already at the first decision of the sequence.

<sup>9</sup> To see how these numbers are derived, consider e.g. the upper left box. If both  $i_2$  and  $i_3$  use 'Trigger' as their strategy, in the first decision they cooperate (logroll) and thus get a payoff of five units each. In the next decision, which occurs with  $p = 0.6$ , they still cooperate (since none defected in the previous round such that the other would cease to logroll). So their expected payoff is  $5p$ . In the third decision, again each cooperates with an expected payoff of

		$i_2$	
		<i>Trigger</i>	<i>Always Defect</i>
$i_3$	<i>Trigger</i>	12.5	7.5
		12.5	1.5
	<i>Always Defect</i>	1.5	2.5
		7.5	2.5

Table 5

So, with this alternative set of assumptions, and in this case, the best strategy for  $i_2$  and  $i_3$  is clearly *Trigger*. This means logrolling is stable.

Much more could be said about the problem of backward induction, and about the additional common knowledge assumption under which cooperating (logrolling) — or defecting — would be best for certain voters.<sup>10</sup> Suffice it here to say that the possibility of logrolling heavily depends on the formulation of this assumption. If some level of uncertainty is introduced (as in the case described in Table 5), logrolling might nevertheless make the logrollers better off and thus be stable.<sup>11</sup> The collective optimality of the weighted majority rule is then once more at peril.

There is another suggestion regarding how to deal with sets of binary decisions on which logrolling might occur. This suggestion is inspired by Arrow.<sup>12</sup> Arrow claims, in response to Buchanan and Tullock's argument that strategic voting may be collectively *beneficial*, that in such cases, the individuation of the options should take this into account. The idea is that voters should vote on complete states of the world, which already comprise the costs and benefits from vote trading.

The suggestion then goes that, in line with Arrow's claim, we should not understand the above *Dinner and a movie* cases as containing two separate binary decisions. Rather, since what is at stake in one of the decisions in a way affects what is at stake in the other for some voters (the logrollers), we must treat them as one compound decision, concerning the set of all possible combinations of the options. More generally, the idea is that we should include all options on which votes might be traded in one single decision. Could this move solve or mitigate the problem that strategic voting may be collectively *damaging*, without relying on the demanding Common Knowledge assumption?

---

$5p^2$ . And so on, up to round  $n$  with an expected payoff of  $5p^n$ . The sum of the expected payoffs from 'Trigger' is thus  $5+5p+5p^2+\dots+5p^n = 5(1+p+p^2+\dots+p^n)$ . Since  $(1+p+p^2+\dots+p^n)$  approaches  $1/(1-p)$  as  $n$  approaches infinity, the sum of payoffs approaches  $5/(1-p) = 12.5$  for  $p = 0.6$ . Similar calculations will generate the values in the other three boxes of Table 5.

<sup>10</sup> For a good overview, see e.g. Skyrms (1998). Cf. even Pettit and Sugden (1989), Sobel (1993) and Bovens (1997).

<sup>11</sup> Cf. Skyrms (1998: 19).

<sup>12</sup> Arrow (1963: 109).

### 7.3.2 Can logrolling be avoided by devising compound decisions?

Let us reconsider the above *Dinner and a movie* cases, concerning the decisions whether to eat *out* or *in* and whether to watch a *comedy* or a *horror* film. There are four possible combinations of options, as illustrated in the next scenario.

***Dinner and a movie (compound).*** Three individuals must decide, as before, whether to eat *in* or *out* tonight and whether to watch a *comedy* or a *horror* film later on. But now, rather than two separate decisions, they face only one decision with four compound options: *out+comedy*, *out+horror*, *in+comedy* and *in+horror*. The case is presented in *Table 6*, where the numbers correspond to the added-up numbers from *Table 2* and *Table 3* above.

	$i_1$	$i_2$	$i_3$
<i>out+comedy</i>	10	1	1
<i>out+horror</i>	5	0	6
<i>in+comedy</i>	5	6	0
<i>in+horror</i>	0	5	5

*Table 6*

The pairwise comparisons of all four options, *out+comedy*, *out+horror*, *in+comedy* and *in+horror*, generate six binary decisions. Non-strategic voting would give us the following results. (The figures in parenthesis specify the numbers of votes either option receives from the respective voter, as indicated by the subscripts. Votes are assigned in proportion to the stakes that can be read off from *Table 6*.)

$$\begin{array}{ll}
 \text{out+comedy} > \text{out+horror} & (5_{i_1} + 1_{i_2} > 5_{i_3}); \\
 \text{out+comedy} > \text{in+comedy} & (5_{i_1} + 1_{i_3} > 5_{i_2}); \\
 \text{out+comedy} > \text{in+horror} & (10_{i_1} > 4_{i_2} + 4_{i_3}); \\
 \text{out+horror} \sim \text{in+comedy} & (6_{i_3} \sim 6_{i_2}; i_1 \text{ has no stakes and thus 0 votes}); \\
 \text{out+horror} > \text{in+horror} & (5_{i_1} + 1_{i_3} > 5_{i_2}); \\
 \text{in+comedy} > \text{in+horror} & (5_{i_1} + 1_{i_2} > 5_{i_3}).
 \end{array}$$

By the Condorcet-style weighted majority rule, we get *out+comedy* > *out+horror* ~ *in+comedy* > *in+horror*. So, as expected, the optimal outcome in terms of collective well-being, *out+comedy*, is declared the winner.

Let us now assume that  $i_2$  and  $i_3$  instead decide on the voting strategy ‘demote *out+comedy*’: when possible, they always vote against *out+comedy*; otherwise, they vote according to their stakes. This means that in the first binary decision,  $i_2$  would cast her pivotal vote for *out+horror* instead, and in the second binary decision,  $i_3$  would cast her pivotal vote for *in+comedy* instead. Then, *out+comedy*, previously the top option, would be ranked be-

low these two tying options and we would get  $out+horror \sim in+comedy > out+comedy > in+horror$ . So with this strategy, the logrollers would promote two collectively suboptimal options to tie for the collective top. Clearly then, the weighted majority rule here cannot ensure the collectively optimal outcome. But can it really be in  $i_2$ 's and  $i_3$ 's self-interest to employ this voting strategy if what they get is a top tie of their best and worst options?

The answer to this question depends on how an outcome is selected if there is a tie. Under non-strategic voting, a top tie would mean that both options are (equally) collectively optimal, so either one may be chosen. Thus the group may employ a random tie-breaker, e.g. a coin toss, which selects either option with an equal chance. This proposal is in line with what is commonly suggested in the literature on voting procedures: the set of options constituting a top tie is 'interpreted as the result of a first screening process, after which every alternative in the set, and no other, is still a candidate to be the final choice'. The final choice is then made by 'some random device'.<sup>13</sup>

The question is then whether such an even-chance gamble between  $out+horror$  and  $in+comedy$  is better for  $i_2$  and  $i_3$  than the non-strategic collective top  $out+comedy$ . We can compare these alternative outcomes in terms of expected individual well-being. It can easily be seen that for both  $i_2$  and  $i_3$  the expected well-being from strategic voting is higher than the expected well-being from non-strategic voting. If they vote according to the above strategy, the expected well-being for each of them is  $\frac{1}{2}(6) + \frac{1}{2}(0) = 3$  units, which is more than the one unit from the non-strategic outcome  $out+comedy$ .<sup>14</sup>

Note that under these circumstances,  $i_1$  does not have any opportunity to 'answer' the coalition's strategy. She is outvoted by  $i_2$  and  $i_3$  in all binary decisions, in the sense of having fewer votes than the coalition, except in the decision on  $out+comedy$  vs.  $in+horror$ . Thus, shifting her votes from  $out+comedy$  to  $in+horror$  is  $i_1$ 's only available effective strategy. Yet it would only reverse the collective (strategic) ranking of the two options at the bottom,  $out+comedy$  and  $in+horror$ , and not affect the top tie at all.

This shows that under the Condorcet-style weighted majority rule — in conjunction with an even-chance lottery — there are cases where it is in some voters' self-interest to vote strategically, where doing so may lead to collectively suboptimal outcomes and where the rest of the voters have no opportunity to counteract this. It might be objected that this is just as expected from the general results on strategic voting, which I discussed in the introduction above. As stated, Feldman shows that when top ties are resolved by an even-chance lottery, and voters know this, only a dictatorship or a

---

<sup>13</sup> Barberà (2011: 787, 790).

<sup>14</sup> I have not discussed the voters' attitudes to risk. In general, I assume voters to be risk-neutral. However, the numbers here leave some room to introduce a certain amount of risk-aversion or -inclination.

duumvirate are non-imposed and strategy-proof.<sup>15</sup> However, note that in the above example, the logrollers would benefit even if the lottery would assign unequal chances to the tying options — e.g. assigning  $2/3$  to the lexicographically first option and  $1/3$  to the other — or if the voters falsely believed some such uneven lottery to apply. Then, one of them would expect  $2/3(6)+1/3(0) = 4$  units, and the other  $1/3(6)+2/3(0) = 2$  units, that is, more than the one unit from the non-strategic outcome.

So the vulnerability of the weighted majority rule to strategic voting does not hinge on the precise nature of the tie-breaker and the voters' knowledge thereof. Moreover, this shows that devising compound options does not save the weighted majority rule from the damaging effects of logrolling — contrary to what one might have hoped when faced with Arrow's suggestion.

Interestingly, what *Dinner and a movie (compound)* also shows is that, when the two separate decisions in which  $i_2$  and  $i_3$  cooperate are combined to one decision with four compound options,  $i_2$  and  $i_3$  can no longer achieve their joint favourite — and collectively worst — option *in+horror*. Recall that they could achieve *in* and *horror* in the two separate binary decisions in *Dinner and a movie 2*. Yet in the compound case they have to settle for a gamble between their individual top and bottom options *out+horror* and *in+comedy*. This gives them slightly less expected well-being (but still more than the non-strategic collective top *out+comedy*). This gamble is also expected collectively better than *in+horror*. So logrolling in the compound case becomes just a little less profitable for the logrollers — and a little less damaging regarding collective optimality. Is this a general feature of such compound decisions?

To answer this question, let us first take a look at multi-option decisions in general (be they decisions with compound options or decisions with more than two simple options) and what kinds of results a coalition of logrollers could achieve. Then, we can go back to the specific kind of compound decisions, constructed from pairs of binary decisions on simple options, that the question refers to.

### 7.3.3 Why logrollers at best can achieve cycles or ties

Let us consider the simple case of a decision with more than two options (the exact number is not relevant). We make the assumption that strategic voting is in the self-interest of two voters,  $i_1$  and  $i_2$ . From this assumption we can infer some things about how the options are ranked (non-strategically) by  $i_1$  and  $i_2$  — how the options are  $i_1$ -ranked and  $i_2$ -ranked, respectively — as well as how they are ranked (under non-strategic voting) by the collective — how the options are  $c$ -ranked.

---

<sup>15</sup> Feldman (1977).

To begin with, we can infer that the option(s) at the top of  $i_1$ 's and  $i_2$ 's individual rankings — the  $i_1$ -top and  $i_2$ -top, respectively — differ from the top option of the collective ranking under non-strategic voting — the  $c$ -top (otherwise strategic voting would not be in their self-interest). Moreover we can infer that the  $c$ -top is  $i_1$ -ranked below the  $i_1$ -top and  $i_2$ -ranked below the  $i_2$ -top (as the  $i_1$ - and  $i_2$ -top is the top of the respective ranking). Finally, we can infer that any option that is  $i_1$ -ranked or  $i_2$ -ranked above the  $c$ -top is  $c$ -ranked below the  $c$ -top (as the  $c$ -top is the top of that ranking). The three rankings thus must look as in *Table 7* (where other options may or may not be inserted in between).

<b><math>c</math>-ranking</b>	<b><math>i_1</math>-ranking</b>	<b><math>i_2</math>-ranking</b>
$c$ -top	$i_1$ -top	$i_2$ -top
...	...	...
$i_1$ -top / $i_2$ -top	$c$ -top	$c$ -top
...	...	...
$i_2$ -top / $i_1$ -top	...	...

*Table 7*

Now, there are three different cases we must distinguish here: (i)  $i_1$  and  $i_2$  might have the exact same individual ranking of all the options and thus also have the same top option ( $i_1$ -top =  $i_2$ -top). Their stakes, that is, their well-being differentials between these options, may differ — but that is of no relevance here. (ii)  $i_1$  and  $i_2$  might have different individual rankings of the options, but with the same top option ( $i_1$ -top =  $i_2$ -top). (iii)  $i_1$  and  $i_2$  might have different individual rankings of the options, with different top options ( $i_1$ -top  $\neq$   $i_2$ -top). I shortly consider these three, one at a time.

For my following arguments, I utilise the notion of being *pivotal*. Previously, I stated that a voter is pivotal in a binary decision if and only if, for a given distribution of all the voters' votes among the two options, had this voter voted for the other option, the outcome would have changed to that option. Now since, in multi-option decisions, binary decision do not generate outcomes, but only a ranking (while the outcome is selected from the top of a combined ranking of several binary decisions), this must be rephrased. So let us now say that a voter is pivotal in a binary decision if and only if, for a given distribution of all the voters' votes among the two options, had this voter voted for the other option, the collective ranking would have changed, such that this other option had been promoted (that is, collectively higher-ranked). This implies, for a group of competent and self-interested voters, that a voter  $i$  is pivotal concerning any two options  $x$  and  $y$  only if  $i$ 's ranking of  $x$  and  $y$  (in terms of her self-interest) coincides with the collective (non-strategic) ranking of  $x$  and  $y$ . In other words: if the (non-strategic) rankings of  $i$  and of the collective, concerning  $x$  and  $y$ , do not coincide, then  $i$  cannot

be pivotal concerning these options. We can immediately make sense of this if we consider that, were  $i$  pivotal concerning  $x$  and  $y$ , while the collective ranking of these options was opposed to hers, then as a pivotal voter she *could* change the collective ranking, and as a self-interested voter she *would* change it.

I also utilise the notion of coalitions being *jointly pivotal*. A coalition of voters is jointly pivotal in a binary decision if and only if, for a given distribution of all the voters' votes among the two options, had this coalition voted for the other option, the collective ranking would have changed, such that this other option had been promoted (collectively higher-ranked). The implications for jointly pivotal coalitions of self-interested voters are similar to the above implications for pivotal self-interested individuals. To spell them out: in a group of competent and self-interested voters, a coalition is jointly pivotal concerning any two options,  $x$  and  $y$ , only if *at least one* of the coalition members' ranking of  $x$  and  $y$  coincides with the collective (non-strategic) ranking of  $x$  and  $y$ . In other words: if the (non-strategic) rankings of *none* of the coalition members coincide with the collective, concerning  $x$  and  $y$ , then the coalition cannot be pivotal concerning these options. We can immediately make sense of this if we consider that, were at least one of the coalition members pivotal concerning  $x$  and  $y$ , while the collective ranking of these options was opposed to hers, then as a pivotal coalition member she *could* change the collective ranking. But does it also hold that, as a self-interested coalition member she *would* change it? It does — as long as being a coalition member (and thus profiting from her coalition partners' strategies) is better for her than not being a coalition member. And this we have assumed.

**(i)  $i_1$  and  $i_2$  have the exact same individual ranking of all the options.** This case looks as in *Table 8*.

<b><math>c</math>-ranking</b>	<b><math>i_1</math>-ranking</b>	<b><math>i_2</math>-ranking</b>
$c$ -top	$i_1$ -top (= $i_2$ -top)	$i_1$ -top (= $i_2$ -top)
...	...	...
$i_1$ -top (= $i_2$ -top)	$c$ -top	$c$ -top
...	...	...

*Table 8*

We can now see that the coalition of  $i_1$  and  $i_2$  cannot be jointly pivotal concerning any pair of options consisting of, on the one hand, the  $c$ -top and, on the other, some option which is  $i_1$ - and  $i_2$ -ranked *above* it, since such a pair's  $c$ -ranking is opposed to its  $i_1$ - and  $i_2$ -rankings. And this means that the coalition cannot promote any one of those options — including the  $i_1$ -top (=  $i_2$ -top) — that are better for its members than the  $c$ -top.

From the description of the case it also follows that any option that is  $i_1$ - and  $i_2$ -ranked *below* the  $c$ -top must also be  $c$ -ranked *below* the  $c$ -top. Thus, the  $c$ -ranking of such a pair coincides with its  $i_1$ - and  $i_2$ -rankings. This means that the coalition may be pivotal — that is, we may construct cases where the coalition is pivotal — concerning such a pair. Such cases may be constructed such that the coalition could achieve either a new *strategic collective top* option, or a *strategic top cycle* among several options, including the  $c$ -top. (Such a cycle would then have to be resolved by some cycle-breaker). For the first possibility: say that the coalition could promote an option that is  $i_1$ - and  $i_2$ -ranked below the  $c$ -top, to the effect of making it the new strategic collective top option. Then, the coalition would settle for an option that is worse for any of its members than the  $c$ -top. In the face of this, it would be in all the coalition members' self-interest to vote non-strategically after all. Thus remains only the second possibility: say that the coalition could promote some option(s), which are  $i_1$ - and  $i_2$ -ranked below the  $c$ -top, to the effect of creating a strategic top cycle that includes the  $c$ -top. This may be in all the coalition members' self-interest (resulting in a greater expected gain in well-being), depending on which cycle-breaker is employed, and how the levels of well-being and risk (from the cycle-breaker) play out for them.

To sum up: for any option the coalition ranks higher than the (non-strategic) collective top, the coalition *cannot* promote it above the collective top, because it is not jointly pivotal. Still, the coalition may be able to promote an option it ranks lower than the (sincere) collective top (we may construct such cases). Then, the coalition can either promote such an option to the effect of making it the new strategic collective top. But this is *against the self-interest* of its members because such an option is worse for any of them than the option they get from voting non-strategically. Or the coalition can promote such an option to the effect of creating a strategic top cycle. This *may* be in its members' self-interest, if the levels of well-being and risk from the cycle-breaker play out for them. Thus, a cycle is the best the coalition can hope for.

**(ii)  $i_1$  and  $i_2$  have different individual rankings of the options, but with the same top option.** Basically, the same structure of the rankings (Table 7) and the same reasoning applies here. Let us look at coalition member  $i_1$ : for any option that is  $i_1$ -ranked above the  $c$ -top,  $i_1$  cannot promote it above the  $c$ -top because she is not pivotal. For any option that  $i_1$  ranks lower than the  $c$ -top, she may be able to promote it. That is, we may construct such cases. Then,  $i_1$  can either promote such an option to the effect of creating a *strategic top cycle*. This may be in  $i_1$ 's self-interest if the levels of well-being and risk from the cycle-breaker play out for her. Or  $i_1$  can promote such an option to the effect of making it the new *strategic collective top*. But such an option is worse for her than the option she gets from voting non-strategically. So making it the new strategic collective top is not in  $i_1$ 's self-interest.

Of course, this may still be in her coalition partner  $i_2$ 's self-interest. Could  $i_1$  then not help her coalition partner out and promote this option to be the collective winner? Well, as a self-interested voter,  $i_1$  would not do this: there is nothing to gain. Recall that we have included *all* the simple options on which logrolling may occur into one multi-option compound decision. Hence there are no further decisions on which  $i_1$ 's coalition partner  $i_2$  could 'pay back' the favour. Thus,  $i_1$  will not promote a lower ranked option above the  $c$ -top, to the effect of making it the new strategic collective top.

The same line of reasoning can of course be applied to  $i_2$ . This shows that, again, a cycle is the best either member of the coalition can hope for.

**(iii)  $i_1$  and  $i_2$  have different individual rankings of the options, with different top options.** Let us assume that this case looks as in *Table 9*. (Note that it could also be the case that the  $c$ -ranking of the  $i_1$ -top and the  $i_2$ -top are reversed — or form a tie. This is accounted for in the argument below.)

<b><math>c</math>-ranking</b>	<b><math>i_1</math>-ranking</b>	<b><math>i_2</math>-ranking</b>
$c$ -top		
...	$i_1$ -top	$i_2$ -top
$i_1$ -top	...	...
...	$c$ -top	$c$ -top
$i_2$ -top		

*Table 9*

We can now see that in a binary decision between the  $c$ -top and the  $i_1$ -top,  $i_1$  cannot be pivotal. But let us assume that  $i_2$  is pivotal in this decision. This means that  $i_2$  ranks the  $i_1$ -top below the  $c$ -top. Similarly, in a binary decision between the  $c$ -top and the  $i_2$ -top,  $i_2$  cannot be pivotal. But let us assume that  $i_1$  is pivotal in this decision. This means that  $i_1$  ranks the  $i_2$ -top below the  $c$ -top. According to these assumptions, the case now looks as in *Table 10*.

<b><math>c</math>-ranking</b>	<b><math>i_1</math>-ranking</b>	<b><math>i_2</math>-ranking</b>
$c$ -top	$i_1$ -top	$i_2$ -top
...	...	...
$i_1$ -top	$c$ -top	$c$ -top
...	...	...
$i_2$ -top	$i_2$ -top	$i_1$ -top

*Table 10*

Then,  $i_1$  can promote the  $i_2$ -top, and  $i_2$  can promote the  $i_1$ -top, such that both are (strategically) collectively ranked above the  $c$ -top. However, which of the  $i_1$ -top and  $i_2$ -top could be made the strategic top option, that is, could be

selected as winner? *Table 10* tells us that  $i_2$  cannot be pivotal concerning these two options, since the options'  $c$ -ranking is opposed to their  $i_2$ -ranking. But  $i_1$  may be pivotal — we could construct such a case. However, if  $i_1$  would vote strategically and thus achieve a reversed (strategic) collective ranking of these two options, she would thereby promote an option to be the new strategic collective top that she ranks below the  $c$ -top, which would be selected by non-strategic voting. Logrolling is then not in her self-interest — and thus she will defect. Apprehending this,  $i_2$  will defect as well. Thus, logrolling will collapse in such a case. (The same reasoning applies, mutatis mutandis, if the  $c$ -ranking of the  $i_1$ -top and the  $i_2$ -top were reversed and  $i_1$  were pivotal in the decision between the  $i_1$ -top and the  $i_2$ -top.)

However, it may also be the case that there is a tie between the  $i_1$ -top and the  $i_2$ -top in the  $c$ -ranking, according to *Table 11*.

<b><math>c</math>-ranking</b>	<b><math>i_1</math>-ranking</b>	<b><math>i_2</math>-ranking</b>
$c$ -top	$i_1$ -top	$i_2$ -top
...	...	...
$i_1$ -top $\sim$ $i_2$ -top	$c$ -top	$c$ -top
...	...	...
...	$i_2$ -top	$i_1$ -top

*Table 11*

Then, again,  $i_1$  may be pivotal concerning the pair of options  $c$ -top and  $i_2$ -top, and  $i_2$  may be pivotal concerning the pair  $c$ -top and  $i_1$ -top (we may construct such a case). Together they could thus promote both the  $i_1$ -top and the  $i_2$ -top, such that both are (strategically) collectively ranked above the  $c$ -top. Then, by not resolving the collective tie between the  $i_1$ -top and the  $i_2$ -top (that is, by voting non-strategically in this decision), they would make both options the new *strategic top tie*. This may be in both their self-interest, if the levels of well-being and risk from the tie-breaker play out for them. Thus a top tie is the best either member of the coalition can hope for. (In fact, this is what occurs in the above *Dinner and a movie (compound)*.)

In summary, what has been shown is that logrolling among coalition members can at best achieve a cycle or a top tie, that is, a gamble on sets of such options that are both better and worse for them than (or include) the (non-strategic) collective top. Since my argument in this section does not specify the exact number of options, it holds for all decisions with more than two options. Moreover, the argument can be extended to settings with a coalition with more than two members. If their rankings of the options coincide, they can be considered as one single coalition member, with all their stakes added up. If they have diverging preference rankings, we can extend case (iii) to more than two voters by making the exact same arguments for voters

$i_4, i_5, \dots i_n$ . (Note that such a coalition by assumption must consist of voters whose top option differs from the  $c$ -top.)<sup>16</sup>

Now the next question is whether such a gamble in compound cases must be less profitable for the logrollers — and a little less damaging from a collective welfare perspective — than the logrolling result in the separate binary decisions of which the compound case is comprised.

### 7.3.4 Why logrolling in compound decisions is worse for logrollers and better for all (than in separate binary decisions)

Recall the kind of case where a multi-option decision with four options is created from joining two binary decisions with two simple options each: the compound case. Let us say that, generally speaking, the simple options are  $x$  and  $y$  in the first decision, and  $z$  and  $w$  in the second, such that the compound options are  $xz, xw, yz$  and  $yw$ .

In such a compound case, we have assumed that there is a group of individuals,  $i_1$  and  $i_2$ , who might gain from logrolling. From this it follows that the  $i_1$ - and  $i_2$ -top, respectively, cannot be identical with the  $c$ -top. Moreover, we can infer that the  $i_1$ - and  $i_2$ -top cannot be identical either. This is so since the compound options are combinations of the simple options from a pair of binary decisions on which logrolling can occur. In each binary decision,  $i_1$  and  $i_2$ , respectively, were assumed to have the opportunity to help the other one out by voting with the other and against her own self-interest. Let us say that  $i_1$  has low stakes in the binary decision  $x$  vs.  $y$ , where she prefers  $x$  — yet votes with  $i_2$ . Then, we know that  $i_2$  prefers  $y$ . And  $i_1$  has high stakes in  $z$  vs.  $w$ , where she prefers  $w$  — and where  $i_2$  ‘pays back’ by voting against her preferred option  $z$ . This means that their self-interests are opposed on each binary decision: the best two options in the binary decisions for  $i_1$  are the worst two in the binary decisions for  $i_2$  — and the other way around. Hence, we can infer that the  $i_1$ -top and the  $i_2$ -top among the *compound* options cannot be the same. Rather, the  $i_1$ -top is the worst compound option for  $i_2$ , while the  $i_2$ -top is the worst compound option for  $i_1$ . So the case looks like the ones in *Tables 10* or *11* above.

According to the argument in the previous section, this means that by logrolling,  $i_1$  and  $i_2$  can at best achieve a top tie among their respective best compound option (if the case looks like in *Table 11* — this in turn presupposes that on the collective ranking the  $i_1$ -top and the  $i_2$ -top form a tie). This means that each of them will face a gamble between her best and her worst option.

---

<sup>16</sup> It should be noted that the above arguments do not take recourse to varying voting weights but operate only with the notion of being pivotal. This means that they also cover cases where *simple* majority rule is employed.

Moreover, we can infer from the description of the compound case that there is a compound option that is the *joint logrollers' top* — let us call it the *l-top*. This is the combination of simple options that would be the outcome from successful logrolling within the two binary decisions, on which each logroller loses some, but gains some more. This *l-top* option is then *i*-ranked by each logroller *i* below her *i-top*, but above the *c-top* (otherwise it would not be worth it to engage in logrolling to start with). Moreover, we can infer from the description that the *l-top* must be at the bottom of the collective ranking. This is so since it is the combination of simple options that would be the outcome from successful logrolling within the two binary decisions. Successful logrolling, however, here means that the outcome of each binary decision is changed, such that the collectively worse option is made outcome (instead of the collectively better which would result from non-strategic voting). The combination of both collectively worse options must be the collectively worst option among the compound options.

The ranking of compound options then looks as in *Table 12*. For clarification of the compound options' internal structure, their simple components (the binary decisions' options) are stated in parenthesis.

<b>c-ranking</b>	<b><i>i</i><sub>1</sub>-ranking</b>	<b><i>i</i><sub>2</sub>-ranking</b>
<i>c-top</i> ( <i>xz</i> )	<i>i</i> <sub>1</sub> -top ( <i>xw</i> )	<i>i</i> <sub>2</sub> -top ( <i>yz</i> )
<i>i</i> <sub>1</sub> -top ~ <i>i</i> <sub>2</sub> -top	<i>l-top</i> ( <i>yw</i> )	<i>l-top</i> ( <i>yw</i> )
<i>l-top</i> ( <i>yw</i> )	<i>c-top</i> ( <i>xz</i> )	<i>c-top</i> ( <i>xz</i> )
	<i>i</i> <sub>2</sub> -top ( <i>yz</i> )	<i>i</i> <sub>1</sub> -top ( <i>xw</i> )

*Table 12*

From *Table 12* we can now see that *i*<sub>1</sub> and *i*<sub>2</sub> *cannot* promote the *l-top* above the *c-top*, since none of them is pivotal in this decision (both the *i*<sub>1</sub>-ranking and the *i*<sub>2</sub>-ranking are opposed to the *c*-ranking concerning these options). Thus, logrolling in these compound cases is slightly less damaging for the entire collective than in combined binary cases, since it cannot make the collectively worst option the winner (and since no gamble that also involves collectively better compound options can be worse than the collectively worst option).

Moreover, as stated above, *i*<sub>1</sub> and *i*<sub>2</sub> can at best achieve a gamble between their best and their worst compound options, respectively. If this gamble is an even-chance lottery, their expected welfare is the sum of the welfare from their best and their worst compound options divided by two. This, however, cannot be higher than the (expected) welfare from the *l-top*. To see that this is so, consider *i*<sub>1</sub>'s situation. As stated before, *i*<sub>1</sub>'s stakes (well-being differentials  $W_{i1}$ ) in *x* vs. *y* are lower than in *z* vs. *w*. Thus:

$$(1) W_{i1}(x) - W_{i1}(y) < W_{i1}(w) - W_{i1}(z)$$

Now, assume that, contrary to the just stated hypothesis, the gamble between the  $i_1$ -top and the  $i_2$ -top would be at least as good for  $i_1$  as the  $l$ -top, that is, assume:

$$(2) \frac{1}{2} [W_{i_1}(x) + W_{i_1}(w) + W_{i_1}(y) + W_{i_1}(z)] \geq W_{i_1}(y) + W_{i_1}(w)$$

From this we can make the following transformations:

$$(3) W_{i_1}(x) + W_{i_1}(w) + W_{i_1}(y) + W_{i_1}(z) \geq 2W_{i_1}(y) + 2W_{i_1}(w) \quad \text{[from (2)]}$$

$$(4) W_{i_1}(x) + W_{i_1}(z) \geq W_{i_1}(y) + W_{i_1}(w) \quad \text{[from (3)]}$$

$$(5) W_{i_1}(x) - W_{i_1}(y) \geq W_{i_1}(w) - W_{i_1}(z) \quad \text{[from (4)]}$$

Yet we know that:

$$(1) W_{i_1}(x) - W_{i_1}(y) < W_{i_1}(w) - W_{i_1}(z)$$

We have derived a contradiction from (1) and assumption (2). Thus, (2) is false in the present setting. That is, the gamble between the  $i_1$ -top and the  $i_2$ -top cannot be at least as good for  $i_1$  as the  $l$ -top. The same line of reasoning can be applied for  $i_2$ , mutatis mutandis. This means that logrolling in these compound cases must be less profitable for the logrollers than logrolling in the separate binary decisions.

To sum up, what has been shown is that logrolling in the above type of four-option compound cases is less profitable for the logrollers — and less damaging for the entire collective — than logrolling in the separate binary decisions the compound case is made up of. This speaks in favour of devising compound decisions from binary decisions.

### 7.3.5 Why logrollers in compound four-option decisions do not face the logrollers' dilemma

Recall that we introduced compound multi-option decisions with the ambition that they should comprise *all* the options on which the coalition could trade votes within one and the same decision. This presupposes that there are not indefinitely many such options, since that would make it impossible to make pairwise comparisons of *all* the options, such that the weighted majority rule could not be applied. In other words, this setting works only for finite sequences of binary decisions. For such finite sequences, I showed that (under certain common-knowledge assumptions) the logrollers' dilemma arises for non-compound binary decisions, such that logrolling would collapse. Interestingly, the logrollers' dilemma need not arise for compound decisions

in such a finite sequence of decisions. To see this, consider the following argument, for a compound decision resulting from two binary decisions. As stated before, the logrollers' dilemma has the same structure as the prisoners' dilemma, so the general case looks as in *Table 13*.

$i_2$	$i_1$	<i>Cooperate</i>	<i>Defect</i>
<i>Cooperate</i>		<i>both cooperate</i> (second best for both)	<i>i<sub>2</sub> cooperates</i> (best for $i_1$ , worst for $i_2$ )
<i>Defect</i>		<i>i<sub>1</sub> cooperates</i> (worst for $i_1$ , best for $i_2$ )	<i>none cooperates</i> (bad for both)

*Table 13*

The question is now whether it is possible that the outcomes from each logroller's decision of whether to *cooperate* or *defect* resemble the above structure in all relevant matters. Let us first check whether the result from *none cooperates* is worse for both logrollers than the result from *both cooperate*. If none cooperates, logrolling collapses, and so the *c*-top ( $xz$ ) will be selected. If both cooperate, logrolling proceeds. Then, as we have seen (in 6.3.4 above), the logrollers can at best achieve a gamble between their respective best and worst options, the  $i_1$ -top ( $xw$ ) and the  $i_2$ -top ( $yz$ ). So is this gamble better than the *c*-top?

Let us focus on  $i_1$  first, using the same framework as in 6.3.4 above. Let us now assume, contrary to the structure of the logrollers' dilemma, that the result from *none cooperates* were at least as good as the result from *both cooperate* and see what this implies.

$$(6) W_{i_1}(x) + W_{i_1}(z) \geq \frac{1}{2} [W_{i_1}(x) + W_{i_1}(w) + W_{i_1}(y) + W_{i_1}(z)]$$

$$(7) 2W_{i_1}(x) + 2W_{i_1}(z) \geq W_{i_1}(x) + W_{i_1}(w) + W_{i_1}(y) + W_{i_1}(z)$$

[from (6)]

$$(8) W_{i_1}(x) + W_{i_1}(z) \geq W_{i_1}(w) + W_{i_1}(y)$$

[from (7)]

$$(9) W_{i_1}(x) - W_{i_1}(y) \geq W_{i_1}(w) - W_{i_1}(z)$$

[from (8)]

Yet we know that:

$$(1) W_{i_1}(x) - W_{i_1}(y) < W_{i_1}(w) - W_{i_1}(z)$$

We have derived a contradiction from (1) and assumption (6). Thus, (6) is false. The same reasoning applies, *mutatis mutandis*, for  $i_2$ . This simply means that the outcome of *none cooperates* (the *c*-top) is worse for each logroller than the outcome of *both cooperate* (the gamble between the  $i_1$ -top

and the  $i_2$ -top). So far, the logrollers' dilemma has the same structure as the prisoners' dilemma.

The next question is whether the outcome of  $i_2$  cooperates (while  $i_1$  defects) is best for  $i_1$  and worst for  $i_2$ . Let us check first whether it is best for  $i_1$ . This would mean that it must be better for  $i_1$  than the outcome of *both cooperate* (the gamble between the  $i_1$ -top and the  $i_2$ -top). Thus, assume:

$$(10) [\text{outcome of } i_2 \text{ cooperates}] > \frac{1}{2} [W_{i1}(x) + W_{i1}(w) + W_{i1}(y) + W_{i1}(z)].$$

The question is now what kind of outcome to expect when  $i_2$  cooperates and  $i_1$  defects. There are several possibilities, but also several constraints here. First of all,  $i_1$  cannot promote her best ( $i_1$ -top) and second best ( $l$ -top) options to be the collective winner, although either would be better for her than the gamble. She simply cannot resolve the tie between the  $i_1$ -top and the  $i_2$ -top to her advantage, and she cannot promote the  $l$ -top above the  $i_2$ -top since she is not decisive on that decision (see 6.3.4). Moreover,  $i_1$  does not want to promote her third best ( $c$ -top) or worst ( $i_2$ -top) options to be the collective winner since they are worse for her than the gamble. (That the  $c$ -top is worse is shown in the previous paragraph; the  $i_2$ -top must be still worse since it is  $i_1$ -ranked below the  $c$ -top.). Since there thus is no single compound option for her to promote to be the collective winner,  $i_1$  can at best hope to achieve some tie or cycle that is better for her than the gamble.

Since  $i_1$  cannot resolve the tie between the  $i_1$ -top and the  $i_2$ -top to her advantage, she cannot achieve a cycle that does not include her worst option (the  $i_2$ -top). And since she (alone) is not pivotal concerning her second worst option ( $c$ -top) and any option she ranks higher, she cannot achieve a cycle that does not include her second worst option either. Thus the best she can hope for is a cycle between her best ( $i_1$ -top), second worst ( $c$ -top) and worst options ( $i_2$ -top), or a global cycle between all four options. Would one of them be good enough for her to make defecting worthwhile?

Let us first check for a cycle of  $i_1$ -top,  $c$ -top and  $i_2$ -top, which, we assume, would then be resolved by an even-chance lottery, assigning  $\frac{1}{3}$  of a chance to each option. Would this be better for  $i_1$  than the gamble she could receive from logrolling? Assume that it were better, such that:

$$(10') \frac{1}{3} [W_{i1}(x) + W_{i1}(w) + W_{i1}(x) + W_{i1}(z) + W_{i1}(y) + W_{i1}(z)] > \frac{1}{2} [W_{i1}(x) + W_{i1}(w) + W_{i1}(y) + W_{i1}(z)]$$

$$(11) 4W_{i1}(x) + 2W_{i1}(w) + 4W_{i1}(z) + 2W_{i1}(y) > 3W_{i1}(x) + 3W_{i1}(w) + 3W_{i1}(y) + 3W_{i1}(z)$$

[from (10')]

$$(12) W_{i1}(x) + W_{i1}(z) > W_{i1}(w) + W_{i1}(y)$$

[from (11)]

$$(13) W_{i1}(x) - W_{i1}(y) > W_{i1}(w) - W_{i1}(z)$$

[from (12)]

Yet again we know that:

$$(1) W_{i_l}(x) - W_{i_l}(y) < W_{i_l}(w) - W_{i_l}(z)$$

We have derived a contradiction from (1) and assumption (10'). Thus, (10') is false. The same reasoning applies, mutatis mutandis, for  $i_2$ . This means that a cycle between their best ( $i_l$ -top), second worst ( $c$ -top) and worst options ( $i_2$ -top) cannot be better than the gamble between the  $i_l$ -top and the  $i_2$ -top for either  $i_l$  or  $i_2$ .

Lastly, let us check for a global cycle among all four options, which, we assume, would then be resolved by an even-chance lottery, assigning  $\frac{1}{4}$  of a chance to each option. Would this be better for  $i_l$  than the gamble she could receive from logrolling? Assume that it were better, such that:

$$(10'') \frac{1}{4} [W_{i_l}(x) + W_{i_l}(w) + W_{i_l}(y) + W_{i_l}(z) + W_{i_l}(x) + W_{i_l}(y) + W_{i_l}(z)] > \frac{1}{2} [W_{i_l}(x) + W_{i_l}(w) + W_{i_l}(y) + W_{i_l}(z)]$$

$$(14) 2W_{i_l}(x) + 2W_{i_l}(w) + 2W_{i_l}(y) + 2W_{i_l}(z) > 2W_{i_l}(x) + 2W_{i_l}(w) + 2W_{i_l}(y) + 2W_{i_l}(z)$$

[from (10'')]

As can easily be seen, (14) is false, as no sum of values can be greater than itself. We have derived a false claim from assumption (10''), which thus must be false. The same reasoning applies, mutatis mutandis, for  $i_2$ . This means that a global cycle between all four options cannot be better than the gamble between the  $i_l$ -top and the  $i_2$ -top for either  $i_l$  or  $i_2$ .

Thus, there is no achievable outcome for  $i_l$  or  $i_2$  individually that would make either one's defecting from logrolling worthwhile. This means that the outcomes from each logroller's decision of whether to *cooperate* or *defect* do not fit into the above structure of the prisoners' dilemma. In other words, the logrollers' dilemma *cannot* arise in such compound cases and hence the logrolling coalition will be stable. This then speaks against devising compound decisions in the considered four-option compound cases.

### 7.3.6 Results so far

To sum up: first, logrolling over several binary decisions is not stable for competent self-interested voters, under a strong Common Knowledge assumption. This instability is due to the logrollers' dilemma. However, by introducing some level of uncertainty within the common knowledge assumption, logrolling may be stable — and potentially collectively damaging.

Second, when several binary decisions are combined into compound decisions, by putting together simple options into compound ones, a logrolling coalition cannot achieve anything better than a tie or a cycle. Moreover, considering a special setting of cases with four compound options, which are

derived from two pairs of simple options, logrolling has been shown to be somewhat less profitable (though still profitable) for the logrollers and somewhat less damaging (though still damaging) from a collective welfare perspective than logrolling in the separate binary decisions that make up the compound case. These then are the (slight) advantages from designing a compound decision, which includes all options on which votes can be traded.

However, third, in the considered four-option compound case, the logrolling coalition is stable, while logrolling in the sequence of separate binary decisions may possibly collapse — at least under a rather strong Common Knowledge assumption. Given this assumption, we may then conclude that we better refrain from designing compound decisions and go for potentially logrolling-immune, and thus efficiency-preserving, binary decision-making after all. One worry with this conclusion is that it has only been shown for this special setting and thus is of limited relevance. Therefore, there is room for further research on whether these results could be generalised.

Fourth, a further upshot of the previous discussion is that sticking to binary decision-making has another advantage over compound decision-making. Consider that, while it takes just two binary decisions in *Dinner and a movie 2*, there are six binary decisions required in *Dinner and a movie (compound)*. And that is just a case with four compound options. If there were three binary decisions with all-in-all six simple options to begin with ( $x$  vs.  $y$ ;  $z$  vs.  $w$ ;  $u$  vs.  $v$ ), there would be eight compound options ( $xzu$ ,  $xzv$ ,  $xwu$ ,  $xwv$ ,  $yzu$ ,  $yzv$ ,  $ywu$ ,  $ywv$ ), requiring 28 binary decisions, and so on. Devising compound decisions would thus require incessant voting, increasing the costs of decision-making and thus indirectly reducing the efficiency of the weighted majority rule. Avoiding compound decisions thus appears even more appealing.

However, irrespective of logrolling, there is another problem. There is the possibility that an individual voter realises that some decisions (binary or multi-option) are not made in isolation from certain others and that she can gain from voting strategically *all by herself*. This individual form of strategic voting, by self-interested voters, is considered in the next section.

## 7.4 Individual strategies for self-interested voters

Logrolling, where several voters ‘trade’ votes with each other, is not the only form of strategic voting. There is also the possibility of individual strategic voting. Strategic voting that is practiced by individual voters, instead of coalitions of voters, presupposes the presence of more than two options that the voter perceives to belong to the same larger issue. This is, for instance, the case in *Dinner plans 4*, where the voters face the three options of eating *in*, eating *out* and picking up the phone to *order* some food (see 7.2 above). By applying the Condorcet-style weighted majority rule, three pairwise decisions are generated. An individual voter can then ‘push’ votes between op-

tions in these binary decisions so as to achieve an outcome that is better for her (but possibly collectively suboptimal). Let us consider a three-option case which closely resembles *Dinner plans 4*, but with a different distribution of stakes.

**Dinner plans 5.** A group of three,  $i_1$ ,  $i_2$  and  $i_3$ , wants to decide whether to eat *in* or *out* or to *order* food.  $i_1$  and  $i_2$  both rank  $in > out > order$ , while  $i_3$  ranks  $out > order > in$ . Their stakes are as stated in *Table 14*.

	$i_1$	$i_2$	$i_3$
<i>in</i>	5	5	0
<i>out</i>	1	1	8
<i>order</i>	0	0	5

*Table 14*

Non-strategic voting would yield the following results:

$$\begin{aligned}
 in &\sim out \quad (4_{i_1} + 4_{i_2} \sim 8_{i_3}); \\
 in &> order \quad (5_{i_1} + 5_{i_2} > 5_{i_3}); \\
 out &> order \quad (1_{i_1} + 1_{i_2} + 3_{i_3} > 0).
 \end{aligned}$$

Thus, we get  $in \sim out > order$ , a ranking where the two collectively optimal options tie for the top. The top tie might then be resolved by an even-chance lottery, such that one of the collectively optimal options *in* or *out* is selected.

Now, assume that  $i_3$ , in the decision on *out* and *order*, uses her three votes to vote for *order* instead of *out*. Then, the following results ensue:

$$\begin{aligned}
 in &\sim out \quad (4_{i_1} + 4_{i_2} \sim 8_{i_3}); \\
 in &> order \quad (5_{i_1} + 5_{i_2} > 5_{i_3}); \\
 out &< order \quad (1_{i_1} + 1_{i_2} < 3_{i_3}).
 \end{aligned}$$

Thus, we get  $in > order > out \sim in$ , a global cycle involving all the options. Assuming that cycles are resolved in the same manner as ties — by an even-chance lottery, there is then a 1/3 chance, respectively, that *in*, *out* or *order* is selected. Effectively  $i_3$  has then managed to decrease the odds for her worst option, *in*, to be selected as outcome. At the same time, she has increased the odds for a collectively suboptimal outcome, *order*.

It is in  $i_3$ 's self-interest to employ this strategy, since her expected payoff is  $\frac{1}{3}(8) + \frac{1}{3}(5) + \frac{1}{3}(0) = \frac{13}{3}$  units. This is higher than the expected payoff from non-strategic voting,  $\frac{1}{2}(8) + \frac{1}{2}(0) = 4$  units. Note that  $i_1$  and  $i_2$  have no strategy for ensuring an outcome with higher expected payoffs for them. They have too few votes to change  $out < order$ . Pushing votes in the decision on *in* and *out* would resolve the tie between these two, but not break the global cycle. And pushing votes in the decision on *in* and *order* would break

the cycle, but at the cost of selecting *order* as outcome, with a payoff of zero units for both  $i_1$  and  $i_2$ .

This shows that the present case for the weighted majority rule gets in trouble regardless of whether logrolling is possible since there may be individual strategic voting. And single voters have no incentive to defect from their own strategies and thus do not get themselves into prisoners' dilemmas. Hence, individual strategies are stable. However, it can be shown that any single self-interested individual can at best achieve a cycle by voting strategically.

#### 7.4.1 Why any individual at best can achieve a cycle

The subsequent argument is essentially the same as the one in 7.3.3 above, concerning case (i)  $i_1$  and  $i_2$  have the exact same individual ranking of all the options. Basically, we now treat these two as being numerically identical. I nonetheless spell out the argument for the case where there is only one strategic voter  $i$ , facing more than two options.

For the following argument, I again utilise the notion of being pivotal, which, as stated above, implies that a voter  $i$  is pivotal concerning any two options  $x$  and  $y$  only if  $i$ 's ranking of  $x$  and  $y$  coincides with the collective (non-strategic) ranking of  $x$  and  $y$ . In other words: if the (non-strategic) rankings of  $i$  and the collective, concerning  $x$  and  $y$ , do not coincide, then  $i$  cannot be pivotal concerning these options.

Now let us consider the simple case of a decision with at least three options. We make the assumption that strategic voting may be rational for some voter  $i$ . From this assumption we can infer how some options are ranked by voter  $i$  — how the options are  $i$ -ranked — as well as how they are ranked (under non-strategic voting) by the collective — how the options are  $c$ -ranked. First, we can infer that the top option of  $i$ 's ranking — the  $i$ -top — is different from the top option of the collective ranking under non-strategic voting — the  $c$ -top. Second, the  $c$ -top is  $i$ -ranked below the  $i$ -top. Moreover, we infer that any option that is  $i$ -ranked above the  $c$ -top is  $c$ -ranked below the  $c$ -top. The rankings thus look as in *Table 15* (other options may be inserted in between).

<b><math>c</math>-ranking</b>	<b><math>i</math>-ranking</b>
$c$ -top	$i$ -top
...	...
$i$ -top	$c$ -top
...	...

*Table 15*

From *Table 15* we can see that  $i$  cannot be pivotal concerning any pairs of options consisting of the  $c$ -top and some option that is  $i$ -ranked *above* it since their  $c$ - and  $i$ -rankings are opposed. Thus,  $i$  cannot promote any options that are better for her than the  $c$ -top.

From the above assumption it also follows that any option that is  $i$ -ranked *below* the  $c$ -top must also be  $c$ -ranked below the  $c$ -top. This means that  $i$  may be pivotal (that is, we may construct cases where  $i$  is pivotal) concerning pairs consisting of the  $c$ -top and some option that is  $i$ -ranked below it since their  $c$ - and  $i$ -rankings coincide. Yet by promoting any such lower  $i$ -ranked option, to the effect of making it the new *strategic collective top* option,  $i$  would settle for an option which is worse for her than the  $c$ -top (which would be the result if she voted non-strategically). This would not be in her self-interest.<sup>17</sup> On the other hand,  $i$  may promote some options that are  $i$ -ranked below the  $c$ -top, to the effect of creating a *strategic top cycle* that includes the  $c$ -top. This may be in her self-interest, depending on the payoffs and risk (from the cycle-breaker) that are involved. In fact, *Dinner plans 4* above exemplifies a case where this would be in the strategic voter's self-interest.

To sum up: any option that  $i$  ranks higher than the (non-strategic) collective top,  $i$  *cannot* promote above the collective top because she is not pivotal. Any option which  $i$  ranks lower than the (non-strategic) collective top,  $i$  may be able to promote (we may construct such cases). Then,  $i$  can either promote such an option to the effect of making it the collective winner. But this is *against her self-interest* because such an option is worse for her than the option she gets from voting non-strategically. Or  $i$  can promote such an option to the effect of creating a cycle involving a number of options. An even-chance gamble among the cycling options can never be as expectedly good for her as her best option, but it can be better than the non-strategic collective top. Thus, individual strategic voting *may* be in  $i$ 's self-interest, if the payoffs and risk from the cycle-breaker play out for her.<sup>18</sup> Thus, a single self-interested voter can at best achieve a cycle.

Again, since the exact number of options is not specified, the argument holds for any decision with more than two options. However, one worry with this conclusion is that it has only been shown to hold for cases where there is one single strategic voter. I leave the further question of what happens if

---

<sup>17</sup> For instance, in the simple case where the (non-strategic) collective ranking is  $x > y > z$ , and  $i$ 's ranking is  $y > x > z$ ,  $i$  may be pivotal on  $x$  vs.  $z$  and  $y$  vs.  $z$  (while she cannot be pivotal on  $x$  vs.  $y$ ). By strategically reversing her ranking on these two binary decisions, she would then change the collective ranking to  $z > x > y$ . But then  $z$  would be the outcome, which  $i$  ranks lower than  $x$ . So this would not be in her self-interest.

<sup>18</sup> It should be noted that the above proof does not take recourse to varying voting weights, but operates only with the notion of being pivotal. This means that it can be extended to cover cases where the simple majority rule is employed.

there are two or more independent strategic voters (who do not form a coalition) to future research.

#### 7.4.2 Results so far

What has been shown is that individual strategies from single voters at best can achieve a cycle. However, this also shows that even in the absence of logrolling opportunities, individual strategies constitute a problem for the case for the weighted majority rule. Single voters have no incentive to defect from their own strategies. And there seems to be no way to present the options in the above *Dinner plans 4 (in, out, order)* in such a way that individuals do not perceive them to have some bearing on each other, beyond the binary decisions in which they figure. Clearly, they all pertain to the larger issue of where to eat tonight. Clearly, they are therefore related to each other. And we must assume voters to see that too — at least as long as they are assumed to have some grasp of what they are voting on. And they need such a grasp to be able to vote in their self-interest (in a systematic, non-random way) in the first place. Thus, strategic voting seems inescapable.

### 7.5 Alternative versions of the weighted majority rule

As we have seen, strategic voting under the Condorcet-style weighted majority rule can at best achieve ties or cycles. It may now be suggested that all collectively suboptimal outcomes depend on the use of inappropriate tie- and cycle-breakers. As has been stated, whether voting is in the interest of some individual or coalition is dependent on the involved payoffs — which are given in each case description and thus non-negotiable — and the risks or chances — which are determined by whatever tie- or cycle-breaker we choose.

Above (in 7.3.2) it was suggested that the set of options constituting a top tie should be ‘interpreted as the result of a first screening process, after which every alternative in the set, and no other, is still a candidate to be the final choice’, to be made by ‘some random device’.<sup>19</sup> This suggestion was then carried over to resulting top cycles — or indeed, even global cycles among all the options. But of course, we can think of other tie- and cycle-breakers.

We could opt for a second vote among the final choice set, but this would not help us whenever strategic voting generates global cycles among all the options. Then, the final choice set contains all the options of the initial choice set. A second vote among them would then not resolve the cycle.

---

<sup>19</sup> Barberà (2011: 787, 790).

One could also opt for a second-stage ‘dictator’ (which might for instance be a democratically elected chairperson) who selects the winner from the final choice set. But of course, this would make strategic voting all the more profitable for such a ‘dictator’, or for those who know they would gain from her decision.

Here is another, more promising tie- and cycle-breaker. Whenever a tie or a cycle occurs, select the collectively best option among the tying or cycling options. Unfortunately, this tie-breaker runs into problems in the above *Dinner and a movie (compound)* case (see 7.3.2 above). There voters  $i_2$  and  $i_3$  form a logrolling coalition and achieve  $out+horror \sim in+comedy > out+comedy > in+horror$  as a collective ranking. Thus, the two suboptimal options  $out+horror$  and  $in+comedy$  tie at the top. As can be seen from *Table 6*, both  $out+horror$  and  $in+comedy$  comprise an aggregate of 11 units. But then the suggested tie-breaker cannot choose between the two tying options at the top since both are equally collectively ranked. And even if it could resolve the tie (e.g. by some additional random device), it would select a collectively suboptimal option after all.

One could of course go straight for a fool-proof tie- and cycle-breaker such that: whenever a tie or a cycle occurs, select the collectively best among *all* the options. (There may be several best options, so we still need an extra, e.g. random device, tie-breaker.) This would mean that, in effect, whenever there are cycles or ties (e.g. from strategic voting), we should ignore the results and go straight for collective welfare maximisation instead. A problem with this approach is that it presupposes more knowledge than hitherto required by the case for the weighted majority rule. As stated before (in 3.4 above), the correct distribution of numbers of votes only requires knowledge of individual *stakes* (well-being differentials), not of the *direction* of individual preferences. But knowledge of the latter would be required for the suggested tie-breaker. So this solution thus rests on a more demanding assumption than the rest of the case. (Another problem is that this seems to turn the whole voting business into something of a travesty. Voters would be required to vote — yet their votes would matter only as long as they voted correctly.)

So far then, no satisfactory tie- and cycle-breaker has emerged. Yet this only means that there is room and a need for further research on how to improve the Condorcet-style weighted majority rule.<sup>20</sup>

---

<sup>20</sup> I have just recently come across an article by Young (1988) in which he discusses Condorcet's suggestion for a cycle-breaker, which may be employed in cases when cycling is due to *uncertainty*. Condorcet's proposal is to consider the pairwise collective rankings of all the options and, if they form a cycle, delete the ranking with the smallest vote differential between the options. Young criticises and improves this proposal, stating that one should consider the pairwise collective rankings of all the options and, if they form a cycle, reverse the ranking within that *set* of pairwise comparisons which has the smallest *sum* of vote differentials between the options, such that the cycle is resolved. This cycle-breaker could possibly be

It could now also be suggested that the problems from strategic voting that threaten the case for the weighted majority rule have been shown to befall the Condorcet-style extension of the rule. Recall that this extension, in form of an *ordinal ranking rule*, was introduced in order for the weighted majority rule to handle multi-option cases. So another field for future investigation concerns the question of whether this ranking rule could be replaced by some other and whether such an alternative extension of the weighted majority rule would be less vulnerable to the above-mentioned problems.

As pointed out before, Brighthouse and Fleurbaey are ambiguous on the point of how multiple options should be ranked under the weighted majority rule. They simply claim, ‘the options are ranked by application of the weighted majority rule over every pair of options’ and subsequently refer to a ‘*numerical ranking*’ of the options, without specifying how such numerical ranking should proceed.<sup>21</sup> Such a specification is crucial, however, since the outcomes may change according to it.

This section thus concludes with the observation that the weighted majority rule needs further refinement, both regarding the tie- and cycle-breakers it might employ and regarding its how it is supposed to work in multi-option settings.

## 7.6 Common-interested voters and strategic voting

Until now we have considered self-interested voters. Are there analogous arguments for common-interested voters when it comes to strategic voting?

Let us take a look at a group, consisting of common-interested voters only, facing a multi-option decision. Assume that these voters satisfy Self- Or Common-Interested Voting, Competence and Success — and that all know that they are common-interested voters who satisfy these assumptions, and that all know that they know, and so on. Then, no voter has an incentive to vote against the collectively better option in any of the binary decisions. No one can ensure a better result than the collectively optimal that results from competent common-interested voting by everyone, and everyone knows that everyone knows this, and so on. So, with the following common-knowledge assumption, strategic voting is ruled out for competent common-interested voters.

---

employed even when cycling is due to *strategic* voting. However, I have not had time to consider this possibility in detail, and it remains a question for further research.

<sup>21</sup> Brighthouse and Fleurbaey (2010: 143f., my italics). According to Fleurbaey, the reference to ‘numerical ranking’ is simply meant to indicate that a ranking of the options by the weighted majority rule (employing the simple ranking rule) under non-strategic voting is *equivalent* to a ranking of the options by comparing their utility numbers (personal communication, May 2009).

**Weaker Common Knowledge:** Every voter knows, and knows that everyone knows, and so on, that every voter satisfies Self- Or Common-Interested Voting — by being common-interested — as well as Competence and Success.

This assumption is weaker than Common Knowledge (see 7.3.1 above), as it does not require knowledge concerning the upper limit or end of the sequence of decisions to be made.

The conclusion that strategic voting is ruled out under these circumstances is in line with an interesting result from the literature on strategic voting. This result is that there *are* non-dictatorial, predictable voting rules that are strategy-proof — if we assume individual preference rankings to be ‘single-peaked’ on a single dimension. In order to explain what it means for preference rankings to be ‘single-peaked’, we first have to establish that there are contexts in which all the options in a given decision can be linearly ordered along a single dimension. This is often illustrated by reference to those political contexts where all parties or policies (the options) can be ordered along a single ‘left–right’ dimension. Presupposing that there is such a single dimension, we can say that one option,  $y$ , is ‘between’ two others,  $x$  and  $z$ , if  $x$  is to the left of  $y$  and  $z$  to its right. This implies that  $y$  is ‘closer’ to either  $x$  or  $z$  than these two are to each other.

In such a context, voter  $i$ 's preference ranking is *single peaked* if and only if (1)  $i$  has a single top-ranked option  $x$ , and (2) if  $y$  is between  $x$  and  $z$ , then  $y$  is preferred to  $z$  by  $i$ . This simply means that if an option ( $y$ ) is closer to  $i$ 's single top ( $x$ ) then another alternative ( $z$ ) is also preferred to this other alternative ( $z$ ).<sup>22</sup>

To illustrate: say there are three options, a *left-wing*, a *centre* and a *right-wing* party. And say that they can be ordered on a single ‘left–right’ dimension such that *centre* is between *left-wing* and *right-wing*. Then, for a voter  $i$  who ranks  $left-wing > right-wing > centre$ , it is *not* the case that her preference ranking is single-peaked. This is so because her ranking violates condition (2). On the single dimension, *centre* is between  $i$ 's single preferred option *left-wing* and the other, *right-wing* — yet  $i$  does not prefer *centre* to *right-wing*. Or in other words: on the single dimension, *centre* is closer to  $i$ 's preferred *left-wing* than *right-wing* is — yet  $i$  does not prefer *centre* to *right-wing*.

What does this have to do with common-interested voters? Well, first of all, we have assumed that it is possible to rank the options in a given decision according to their collective goodness. (This is the *independent standard* referred to in Chapter 5.) This means that we can linearly order them along a single dimension, from collectively worst to collectively best. Now, assume that a common-interested voter  $i$  is fully competent. Then, she al-

---

<sup>22</sup> Cf. Barberà (2011: 759). For an extensive treatment of single-peaked preferences and their relevance for strategy-proof voting rules, see even Black (1958).

ways correctly perceives what is in the common interest. Assume moreover that she ranks collectively better options above the collectively worse ones. Then her preference ranking must be single-peaked along the dimension of collective goodness.

To see that this is so, consider any three options  $x$ ,  $y$  and  $z$ . Say that voter  $i$  top-ranks  $x$ . Since  $i$  is assumed to be fully competent and common-interested, this means that  $x$  is collectively best. Now, on the collective-goodness dimension, let us say that  $y$  is between  $x$  and  $z$ . Then, we can infer that  $i$  must rank  $y$  over  $z$ . This is so since ‘being between two options’ here means ‘being collectively worse than one (this must here be  $x$ ) and collectively better than the other (this is thus  $z$ )’. Being fully competent and common-interested,  $i$  can never rank this ‘in-between’ option  $y$  below the collectively worse option  $z$ . Hence, she cannot violate conditions (1) and (2). So her preference ranking must be single-peaked.

(A short aside: the link between the voter's individual ranking of the options and the dimension of collective goodness is broken if we assume competent *self*-interested voters. The reason is that they rank the options not according to the options' levels of collective goodness, but according to their levels of individual goodness. Of course, for every voter  $i$ , we could construct an individual dimension on which her rankings are single-peaked — e.g. the total amount of welfare they comprise for  $i$ . But these individual dimensions may turn out to be irreconcilable. That is, it may turn out that there is no way to find a *single* dimension on which *all* self-interested voters' preference rankings are single-peaked. In fact, this is the problem that underlies, e.g. *Dinner plans 3*, as described in Chapter 1, where  $i_1$ 's ranks  $in > out > order$ ;  $i_2$  ranks  $out > order > in$ ; while  $i_3$ 's ranks  $order > in > out$ . There is no single dimension on which all three rankings can be single-peaked.)

What if the above-mentioned strong common-knowledge assumption does not hold? That is, what if some voter  $i$ , while satisfying Self- Or Common-Interested Voting, Competence and Success, were uncertain about the other voters' satisfying these assumptions? Now, it might seem that this would not even be possible. Given this uncertainty, one might want to argue,  $i$  would hardly be able to *correctly* judge which way of voting promotes the common interest, so she would not satisfy Competence, although this was assumed. However, even given complete ignorance about the other voters' voting motivation, a voter might still correctly judge her best strategy, even though she could not infer it from the others' strategies. Recall that ‘correctly judges’ does not require knowledge or justified belief — for which evidential and inferential circumstances would matter. So Competence is not as demanding as to imply certainty about other voters' voting motivation.

However, given first- or higher level uncertainty about other voters' motivation, voter  $i$  could suspect all kinds of erratic or strategic voting behaviour from others and then set out to rectify the result by adopting a counter-strategy of her own. That is, without the Weaker Common Knowledge as-

sumption, strategic voting might be possible — and possibly collectively damaging — even for common-interested voters. To be sure, all individual rankings of the options would still be single-peaked along the dimension of collective goodness, but voters may not vote according to these individual rankings.

Thus, again, given a certain level of common knowledge, common-interested, competent and successful voters would vote non-strategically. Yet given uncertainty about the others' strategies, strategic voting may reappear even for common-interested voters.

## 7.7 Less than fully competent voters

All results in this chapter rest on the demanding assumption of full voter competence when it comes to assessing the options in the light of the voter's self- or the common interest. That is, I have presupposed that voters satisfy the strong Competence assumption. However, consider what happens if we relax this assumption (as was done in Chapter 5).

Recall that for pairwise decisions, the *fifth argument from weak collective optimality* gave us the following conclusion (see 5.2.7 above).

(6) For binary decisions with exactly one common-interest option (according to the given criterion of the common good) and set of a sufficiently large number  $n = 2m+1$  of on average tolerably correlated votes cast by any number  $l < n$  voters, given Minimal Average CJT-Competence and Voting According to Judgment, the weighted majority rule is weakly collectively optimal (according to this criterion).

Now, the phrase ‘is weakly collectively optimal’ was meant to be interpreted as ‘selects the common-interest option with near certainty’. In the present multi-option context, we can then interpret ‘select’ as simply referring to ‘rank higher’. The common-interest option, in turn, is simply the collectively better option (among the two at hand). Thus, the argument gives us the conclusion that for every pairwise decision, if there is a sufficiently large (odd) number of on average tolerably correlated votes cast by some smaller number of voters, if the voters are minimally CJT-competent on average and vote according to their judgment, the weighted majority rule is nearly certain to rank the collectively better option higher, among the two.

If there are more than two options, I stated that the Condorcet-style weighted majority rule would generate pairwise rankings of all the options, derive a collective ranking from these and select the top-ranked option as the outcome. The probability that this ranking is the correct ranking (corresponds to the options' ranking in terms of collective optimality) is then the product of the probability for each pairwise collective ranking to be correct. This means that if there are some number  $o$  of options — generating

$\frac{1}{2}o(o-1)$  of pairwise rankings all of which are correct with probability  $P_n$  — then the collective ranking of all options is correct with probability  $P_n^{\frac{1}{2}o(o-1)}$ . So, even if  $P_n$  is sufficiently high to be called ‘nearly certain’, given a large number  $o$  of options, the collective ranking will be correct with a much lower probability.

To illustrate this problem: consider a case where there are five options, and thus 10 possible pairs of options on which to vote. If the probability of a correct majority vote  $P_n$  is 0.9 for every pairwise decision, the probability of a correct collective ranking of all the options is only  $0.9^{10} = 0.35$ . If the former probability is 0.95, the latter is still only 0.6. If the former is 0.99, the latter is 0.9.

And this is a very modest multi-option case with only five options. Consider a case with 10 options, and thus 45 possible pairs. Then the probability of a correct collective ranking is 0.009, 0.099 and 0.64 for pairwise rankings with  $P_n = 0.9$ , 0.95 and 0.99, respectively. These illustrations suggest that group competence on pairwise rankings must be extremely high — or the number of options extremely small — in order for the results from the former chapter to carry over to multi-option contexts.

Of course, what matters for the collective optimality is not the entire collective ranking for all the options, but only its top-ranked option. The probability of a correct collective top will be higher than probability of a correct entire collective ranking. Still, even if we only consider top-ranked options — in order for the collectively optimal option(s) to be top-ranked, they must not be ranked below — or cycle with — any of the other, collectively suboptimal options. Assume that there are five options, of which exactly one is collectively optimal. Then, this option enters into four pairwise decisions with the other ones — and in all four, it must beat its rival. That is, in these four pairwise decisions, the collective ranking must be correct. If the probability of a correct collective ranking is 0.9 for each pairwise decision, the probability that the collectively optimal option beats each of them is  $0.9^4 = 0.66$ . If the former probability is 0.95 (or 0.99), the latter is 0.81 (or 0.96). If there are 10 options, the collectively best option must be correctly ranked above the nine others. This probability is 0.39, given a 0.9 probability of correct pairwise rankings, or 0.63 (or 0.91) given a 0.95 (or 0.99) probability of correct pairwise rankings.

Thus, the risk of an incorrect outcome rises considerably in multi-option decisions, once the Competence assumption is relaxed. It increases with decreasing group competence on pairwise decisions and with increasing numbers of options. This also means that cycles in the collective ranking can no longer be ruled out, even in the absence of strategic voting.<sup>23</sup> Voter uncertainty may thus undermine the collective optimality in multi-option settings.

---

<sup>23</sup> For less than maximally (CJT-)competent voters, there is an interesting result by Young (1988: 1239), comparing two extended versions of simple majority rule that may be applied in

## 7.8 Conclusions

In this chapter, I have shown that the case for the weighted majority rule may be extended from a binary to a multi-option decision context. The *further-extended argument from collective optimality* showed that, given the Self- Or Common-Interested Voting, the Condorcet-style weighted majority rule selects the common-interest option among multiple options.

However, I then argued that this result hinges on the exact interpretation of Self- Or Common-Interested Voting, namely as Local-Scope SCI-Voting. Under this interpretation, the assumption applies to each pairwise decision considered in isolation. When interpreted as Global-Scope SCI-Voting, applying to an entire multi-option decision, the case for the weighted majority rule is undermined by the possibility of strategic voting. Global-Scope SCI-Voting is implied by the conjunction of Self- Or Common-Interested Voting, Competence (understood as requiring correct judgments that are not myopically limited to the options but concern the available ways of voting) and Success.

Taking the latter three assumptions for granted, and starting out with competent self-interested voters, I first analysed logrolling. I concluded that logrolling is not stable given the following assumption.

**Common Knowledge:** Every voter knows, and knows that everyone knows, and so on, (a) about the upper limit — or end — of the sequence of decisions they face and (b) about every voter's satisfying Self- Or Common-Interested Voting — by being self-interested — as well as Competence and Success.

Under this assumption, the voters face the logrollers' dilemma. However, given a weaker assumption, allowing uncertainty regarding the sequence of decisions faced by all, strategic voting has been shown to be possible and collectively damaging.

I also considered an alternative proposal to deal with sets of several decisions with only two options each: to combine these into one single decision with compound options. I showed that strategic voting on such compound

---

multi-option decisions: the *Condorcet rule* (for each possible pair of the options, each voter cast her vote for one of them; option  $x$  is collectively ranked above option  $y$  if and only if  $x$  receives more votes than  $y$  in the pairwise vote; the collectively top-ranked option is then selected) and the *Borda rule* (each voter submits a complete linear ranking of the options; on each ballot, the lowest-ranked option receives 0 points, the next lowest 1 point, and so on; the options are collectively ranked according to the sum of their points across all ballots; the collectively top-ranked option is selected). Young shows that 'the Borda winner is in fact a better estimate of the [correct option] provided that  $[c]$  is close to  $\frac{1}{2}$ ' and that 'if  $[c]$  is not close to  $\frac{1}{2}$ , then it is still very likely that the Borda winner is the [correct option], even though strictly speaking it may not be the optimum estimate of the [correct option]'. This calls for a closer examination of whether it would be possible to devise a 'Borda-style' version of the weighted majority rule and for an assessment of its performance in multi-option contexts. I leave this question to be dealt with another day.

decisions is less profitable for the logrollers and less collectively damaging. This can be taken to speak in favour of devising compound options. However, doing so may help the logrollers to avoid the logrollers' dilemma, thus making logrolling stable (even given Common Knowledge). Another worry with devising compound options stemmed from the fact that it would result in an exponential increase in pairwise decisions and thus require incessant decision-making. Moreover, as the number of options increases, the probability of a correct collective ranking — or of a correct collective top — decreases, as argued in 7.7 above. Recall also that when voters are less than fully competent, creating compound options may decrease voter competence, as argued in 6.4 above. These results then provide additional reasons against devising compound options.

Second, I analysed strategic voting when practiced by individual, competent and self-interested voters. I argued that they can at best achieve a cycle. This means that they cannot make sure their best option wins, but must settle for a gamble. Still, this gamble may give them greater expected payoffs than the non-strategic outcome. Thus, individual strategic voting is possible — and potentially collectively damaging. Moreover, individual voters do not face the logrollers' dilemma. Individual strategic voting is thus stable.

The observation that logrollers and individual strategic voters can at best achieve ties and cycles brings out the need to focus on the involved tie- and cycle-breakers. One question I identified for further research is thus whether the weighted majority rule could be equipped with a tie- and cycle-breaker that makes strategic voting less or not at all profitable. Another further question is whether there is an alternative and better way of extending the weighted majority rule to multi-option settings, than the here proposed Condorcet-style interpretation.

Third, I considered competent common-interested voters. I argued that given the following Weaker Common Knowledge assumption, they vote non-strategically.

**Weaker Common Knowledge:** Every voter knows, and knows that everyone knows, and so on, that every voter satisfies Self- Or Common-Interested Voting — by being common-interested — as well as Competence and Success.

However, under uncertainty about the others' strategies, the possibility of strategic voting resurfaces.

Finally, I analysed the effects of relaxing the Competence assumption. That is, I once again considered less than fully competent voters. I argued that in in multi-option decisions, even with minimal average voter competence, the risk of an incorrect outcome (or cycle) increases with decreasing group competence on pairwise decisions and with increasing numbers of options. Thus, if there are many options on the table, group competence on pairwise decisions must be extremely high, for collectively optimal out-

comes to be ‘nearly certain’ when Competence is relaxed. An extremely high group competence can be achieved in two major ways: by there being an extremely large number of sufficiently uncorrelated votes, or by there being extremely (if not fully) competent voters. For moderately sized groups who face a many-option decision, this means that Competence cannot be relaxed very much.

With all this said, there are of course other ways to rescue the case for the weighted majority rule for multi-option decisions. Recall that the *further-extended argument from welfare efficiency* goes through if Local-scope SCI-Voting is assumed. Local-scope SCI-Voting is inconsistent with the conjunction of Self- Or Common-Interested Voting, Competence (in a non-myopic sense) and Success. Yet it should be pointed out that Local-scope SCI-Voting is implied by alternative assumptions about the voters' motivational set-up. Consider, for instance, the following.

***Expressivist Motivation:*** Every voter is an expressive voter, in the sense that she desires to express her self-interest or the common-interest when voting.<sup>24</sup>

***Expressivist Competence:*** Every voter has a correct belief that voting for some option  $x$ , among the given alternatives, expresses her self-interest or the common-interest.

***Expressivist Success:*** Every voter acts according to the belief-desire pair referred to in Expressivist Motivation and Expressivist Competence, respectively.

Alternatively, one could reformulate these assumptions for honest — rather than expressive — voters, or the like. Such voters would then, in Dummett's initially quoted words, view the voting paper as a questionnaire, rather than as an instrumental part of a collective decision-mechanism, which generates outcomes that matter to them in their own rights. This is, however, not an approach I consider any further in this study.

---

<sup>24</sup> On the ‘expressive theory of voting’, see Brennan and Lomasky (1993).



## 8 Summary and discussion

### 8.1 Introduction

In this study I have set out to analyse the performance of a novel rule of democratic decision-making: the weighted majority rule. This rule is quite similar to the well-known simple majority rule, but instead of assigning one vote to every voter for each decision, it assigns numbers of votes (as indivisible vote bundles) in proportion to the voters' stakes in the decision. That is, someone who has greater stakes (in the properly defined sense) in a decision receives more votes than someone who has small stakes, whereas an individual who has no stakes at all in this decision does not receive any votes either. I have defined the weighted majority rule as follows.

***The Weighted Majority Rule:*** For all individuals and any decision with two options, (a) every individual is assigned a number of votes in proportion to her stakes, and (b) the option that receives a majority of votes is selected as outcome.

My study started out from an argument for the weighted majority rule in terms of its collective optimality. The *original argument from collective optimality*, as stated by Brighouse and Fleurbaey and separately by Fleurbaey, shows that this rule selects the collectively best of any two given options, if voters vote according to their self-interest. The authors propose that 'collectively best' can be understood in accordance with a sum-total criterion (maximising the sum-total of individual well-being) or with a prioritarian criterion (maximising the sum of individual well-being while giving priority to worse off individuals).<sup>1</sup>

The main result of my study is that this *original argument* can be considerably improved, in the sense that the conditions on which it rests can be relaxed, along the following three main lines. (1) The *original argument* can be adapted to other criteria of the common good. (2) The assumption of self-interested voting can be logically weakened (e.g. to allow voting in the common-interest as well). (3) The scope of the argument can be extended to

---

<sup>1</sup> Brighouse and Fleurbaey (2010), Fleurbaey (*mimeo*).

decisions with more than two options. In the following three sections, I summarise my arguments for these three claims and point out the assumptions on which they rest. (References to the relevant sections are stated in the footnotes.) I critically assess the relevance of my findings and occasionally illustrate them with some of the decision-making cases described throughout this study. Section 8.5 concludes.

## 8.2 The original argument can be adapted to other criteria of the common good

In Chapter 2 I have proposed that ‘the common good’ of a group (or ‘what is in the common interest’) can be understood as some aggregate of the levels of well-being of the individuals within the group. There are numerous suggestions about the kind of aggregation that is required. I have stated a sum-total criterion of the common good, along with prioritarian and sufficientarian criteria, maximin and leximin criteria, and criteria that aggregate individual well-being under certain (non-welfarist) constraints, e.g. concerning the basic rights, autonomy, liberty or desert of the individuals. I have moreover stated a number of theories concerning how individual well-being (what is good for individuals or what is in their self-interest) could be understood. These theories can be summarised under the headings of hedonist, desire-fulfillment and objective list theories.<sup>2</sup>

### 8.2.1 The original argument from collective optimality

In Chapter 3 I showed that this *original argument* can be adapted to fit any combination of these theories of the common good and of individual well-being. That is, I showed that, for any chosen theory of individual well-being (within certain stated measurability and comparability constraints), the argument can be made from all the proposed criteria of the common good. The adaptation of the argument was done by varying the definitions of ‘stakes’, in proportion to which the weighted majority assigns numbers of votes. For instance, when ‘stakes’ are defined as ‘differentials in well-being between the options’, the weighted majority rule is shown to select the collectively best option according to the sum-total criterion of the common good. If, on the other hand, ‘stakes’ are defined as ‘well-being differentials below a given sufficientarian threshold’, the weighted majority rule can be shown to select the collectively best option according to a sufficientarian criterion of the common good. And again, for a maximin criterion of the common good, we identify as stake-holder that individual who among all given options is

---

<sup>2</sup> See 2.3 above.

made worst off.<sup>3</sup> The upshot of my arguments for the adaptability of the original argument was the claim that the argument can be restated in general terms as follows.

### 8.2.2 The generic argument from collective optimality

My *generic argument* rests on the premise (among others) that ‘stakes’ are appropriately defined, according to the given criterion of the common good, and concludes that the weighted majority rule is collectively optimal for any of the proposed criteria.<sup>4</sup> This means that the weighted majority rule can be shown to be optimal for a larger class of normative criteria than previously claimed. The argument is not conditional upon any specific criterion of what is collectively optimal (or, for that matter, any specific theory of individual well-being). The upshot of my results is that the argument from collective optimality, and thereby the weighted majority rule as a method of collective decision-making, is shown to be relevant for a larger domain of moral philosophy.

This relevance is further increased by the observation that, contrary to what many seem to believe, the application of the weighted majority rule is informationally less demanding than the direct application of the relevant criterion of the common good, when the collectively best option is to be identified. To see this, consider a sum-total criterion. In order to determine the individual stakes in a decision and to assign stake-proportional votes, we need not know how the individuals rank the options (whether they rank  $x$  above  $y$  or vice versa), but only the size of their well-being differential between the options. This informational advantage has been argued to hold for all the discussed criteria except, occasionally, maximin and leximin. (For the latter two, if, among all the options, more than one individual is made worst off, we need to know these individuals' rankings of the options to know whom to identify as stake-holder.)<sup>5</sup>

One might now speculate about how accommodating different criteria of the common good (defined on different theories of individual well-being), would affect the practical relevance of the argument. Recall that relaxing the conditions on which the argument is stated — such as allowing a larger class of criteria of the common good — need not make the argument more practically relevant. After all, it might be the case that none of the proposed criteria is ever practically relevant. However, this is clearly an overly pessimistic proposal. Surely there are contexts of decision-making in which we seek to maximise e.g. the sum-total of individual well-being (in some sense).

---

<sup>3</sup> See 3.2 above.

<sup>4</sup> See 3.3 above. Presumably, this might hold even for criteria that I have not discussed. Yet this is not shown by the arguments in this study and thus pending on further investigation.

<sup>5</sup> See 3.4 above.

We might suspect this to be the case in the decision-making of the EU Council of Ministers, where each of the 27 European member states has one seat in the Council, yet the voting weights for member states differ, being in (rough) proportion to their numbers of citizens.<sup>6</sup> If we assume that each citizen matters equally, such that each is taken to have an equal stake in a binary decision at hand, assigning citizen-proportional votes to the representatives and selecting the option that receives a majority of votes amounts to selecting the option which makes more people better off. This can thus be claimed to serve the end of maximising the sum-total of well-being, thus understood, or to formulate it in a less controversial way, of making as many citizens as possible better off. If, however, the EU Council would want to settle for another end, such as maximising the sum-total of well-being for all those below a sufficientarian threshold, the arguments of my study can show a way to achieve this end instead.

To sketch another example, consider weighted voting among shareholders of a company. Within this context, maximising the sum-total of their financial gains certainly appears to be a conceivable end, and financial gains are conceivably taken to be a measure of what makes the share-holders better off. My *generic argument from collective optimality* is then of obvious relevance for these real-life decision-making contexts.

Speaking of practical relevance, there might be a nagging worry concerning whether it could ever be possible to correctly assess peoples stakes in a decision. Would not the purported practical impossibility of correct stakes-assessment make the weighted majority rule inapplicable and thus render my argument practically irrelevant? As should be clear from the just stated pair of real-life examples, this worry is exaggerated. Clearly, there are cases where the stakes can be accurately assessed (e.g., as the number of represented citizens or as the number of owned shares). Moreover, I have in Chapter 3 referred to a proposal stating that, even when there is uncertainty about the stakes, by making the weighted majority rule operate on expected stakes, it can be shown to be *expectedly* collectively optimal.<sup>7</sup>

Regarding the practical relevance of my results, then, I want to maintain that, while it is implausible to claim that the weighted majority rule is easily applicable in all collective decision-making contexts, it is equally implausible to claim that there are no decisions where this rule may be applied. The applicability of the rule and, as an upshot, the relevance of my arguments will depend, among other things, on the nature of the decision and on the presupposed criterion of the common good and theory of well-being.

---

<sup>6</sup> Cf. 1.2 above. The most populated states (Germany, France, UK, Italy) are currently assigned 29 votes each, while the least populated (Malta) has only three, with the other states ranging in between.

<sup>7</sup> See 3.4 above.

## 8.3 The assumption of self-interested voting can be logically weakened

The *generic argument from collective optimality* was the starting point for the following line of inquiry. Just like the *original argument*, it has been stated on the following assumption.

***Self-Interested Voting:*** Every voter (that is, individual who has been assigned a positive number of votes) votes according to her self-interest.

The main part of my study has been dedicated to rendering this assumption logically weaker.

### 8.3.1 The extended argument from collective optimality

In Chapter 4 I have stated an argument that shows that Self-Interested Voting can be relaxed to allow even common-interested voting. The argument started out from a case where all voters vote exclusively according to their self-interest, for which the *generic argument* has shown the weighted majority rule to be collectively optimal. Then, I replaced one of these self-interest voters with one who votes according to the common interest. This was shown not to affect the collective optimality of the weighted majority rule. By successively replacing two, three (etc.) self-interest voters with common-interest voters, while showing the same conclusion to hold, I finally arrived at a case where all the voters vote exclusively according to the common interest. For all these cases, the weighted majority rule was shown to select the collectively best option. This stepwise argument I called the *extended argument from collective optimality*. It was stated on the following assumption, which is logically weaker than (that is, which is implied by, but does not imply) the Self-Interested Voting assumption.

***Self- Or Common-Interested Voting:*** Every voter votes according to her self-interest or according to the common interest.

I also argued that this assumption cannot be further relaxed to allow voting according to other interests (such as the partial interest of a proper subgroup of the entire group, e.g. one's family), by showing how on such an assumption the weighted majority rule might select collectively sub-optimal outcomes.<sup>8</sup>

Self- Or Common-Interested Voting still seems to be a quite strong assumption. However, it should be noted that, as an assumption of the *extended argument*, it is not a necessary condition for the collective optimality of the weighted majority rule, but merely jointly sufficient together with the

---

<sup>8</sup> See 4.2 above.

other conditions (of a binary decision with appropriate stake-assignment and application of the weighted majority rule). Indeed, I argued that this assumption can be limited to the group of pivotal voters. A voter is *pivotal* in a binary decision if and only if, for a given distribution of all the voters' votes between the two options, had this voter voted for the other option, the outcome would have changed to that option. Now the assumption can be further relaxed as follows.

***Pivotal Self- Or Common-Interested Voting:*** Every voter votes according to her self-interest or according to the common interest, or is non-pivotal.

Replacing Self- Or Common-Interested Voting with this new assumption preserves the weighted majority rule's collective optimality. This should come as no surprise, as any non-pivotal voter, according to the stated definition, could not change the outcome from being collectively optimal to another option.<sup>9</sup>

In addition, I argued that only majority stake-holders are in a position to violate Self- Or Common-Interested Voting, that is, to vote against both their self-interest and the common interest in a binary decision. This is so since the majority stake-holders are stipulated to be that group of voters with the same self-interest option who hold the majority of stakes. Thus, I argued, whatever option is in their self-interest, is also in the common interest (according to how the common interest was defined, in terms of an aggregate of the individual interests). If any of them votes against this option, Self- Or Common-Interested Voting is violated. In contrast, all the other voters are minority stake-holders (being that group of voters with the same self-interest option who hold the minority of stakes). For these voters, their self-interest option is opposed to the common-interest option in any binary decision. Thus, whichever option they vote for, it will either be in accordance with their self-interest or with the common interest. Hence we can relax the above assumption further, by limiting it to majority stake-holders rather than the group of all voters.

***Pivotal Non-Erratic Majority Stake-Holders:*** Every pivotal majority stakeholder votes non-erratically, that is, votes according to her self-interest or according to the common interest.

To conclude my discussion of the relaxed Self-Interested Voting assumption: I have shown that this assumption can be replaced by much weaker assumptions without compromising the result of the weighted majority rule's collective optimality. This means that the case for this rule is considerably strengthened. Moreover, it makes the argument potentially more relevant for real-life decision-making. Requiring all voters to vote according to their self-

---

<sup>9</sup> See 4.3 above.

interest is highly unrealistic. Allowing, in addition, common-interested voting makes for a stronger argument. Yet the Self- Or Common-Interested Voting assumption is still overly demanding when it comes to real-life voting. However, requiring such non-erratic voting behaviour only from pivotal majority stake-holders seems in comparison much more modest.

Whether self- or common-interested voting can be expected of real voters is an empirical question, definite answers to which lie beyond the scope of my study and expertise. However, in order to evaluate this issue further I have sketched a possible motivational picture underlying the Self- Or Common-Interested Voting assumption. I suggested that the latter is implied by the conjunction of three assumptions concerning the voters' beliefs and desires.

***Self- Or Common-Interested Motivation:*** Every voter is either self-interested, in the sense that she desires to promote her self-interest, or common-interested, in the sense that she desires to promote the common interest.

***Competence:*** Every voter has a correct belief that voting for some option  $x$ , among the given alternatives, promotes her self-interest (if she is self-interested) or the common-interest (if she is common-interested).

***Success:*** Every voter acts according to the stated belief-desire pair.

These assumptions may replace the Self- Or Common-Interested Voting assumption in the *extended argument from collective optimality*. I have then focused especially on the Competence assumption in an attempt to relax it.

Prior to that, one brief note about the Success assumption: clearly it might not hold in many real-life cases of collective decision-making. Real voters make all kinds of mistakes when setting about to vote, such as checking the wrong box on the ballot. Or they might suffer from weakness of will and other such internal impediments. Or they might face external obstacles, such as coercion. However, we can think of scenarios where voters succeed in voting according to their belief-desire pairs. This is especially likely given certain institutional arrangements that may be brought in place. For instance, if the voting procedure is sufficiently transparent, this may reduce voter mistakes; if the secret ballot is enforced, this may reduce external obstacles. Thus, the Success assumption can give us some guidance on how to improve real-life decisions to ensure the weighted majority rule collective optimality.

### 8.3.2 Arguments from weak collective optimality

These arguments rest on the well-known Condorcet jury theorem (CJT) and a number of extensions and generalisations of it. These Condorcet theorems are devised for showing the conditions under which *simple* majority rule selects the correct outcome, according to an independent standard, such as

truth, when voters judge the options in the light of the independent standard, that is, when voters are truth-trackers. More specifically, the Classical CJT states that, for decisions with two options, if the voters are equally better than chance to judge the options correctly (equally minimally competent), vote according to their judgments and vote independently of each other, the probability that the majority votes for the correct option increases as the number of voters increases, toward certainty as this number increases to infinity. (This last phrase is called the asymptotic conclusion.) Thus for sufficiently large numbers of voters, given equal minimal competence, the majority is nearly certain to vote for the correct option.

My strategy for employing this theorem and various versions of it was to state the conditions each presupposes and then to show that these conditions are satisfied in different cases of decision-making by *weighted* majority rule that I discuss. I first argued that the criterion of the common good (whichever is chosen) can work as an independent standard. Second, I noted that the weighted majority rule is extensionally equivalent to the simple majority rule in all cases where everyone's stakes are equal, where both rules assign one vote to each individual and select the option which gets the majority of votes. I therefore started out by considering only equal-stakes cases. Third, I argued that, even though voters are not assumed to be truth-trackers in this study, when all are common-interested, all judge the options in the light of the independent standard of the common good. I hence started out with common-interested voters.<sup>10</sup>

My first result thus concerned equal-stakes cases with common-interested voters. Employing the classical Condorcet jury theorem, I showed that in these cases, the weighted majority rule is nearly certain to select the common-interest option under the following competence assumption.

***Equal Minimal Competence:*** Every voter is equally likely to judge the options correctly and somewhat more likely than not to judge the options correctly, such that every voter's individual competence is  $c > \frac{1}{2}$ .

I stated that 'being nearly certain to select the common-interest option' is to be weakly collectively optimal. Thus, the weighted majority rule was shown to be weakly collectively optimal even given a much weaker competence assumption — given a sufficiently large number of voters.<sup>11</sup>

I then introduced self-interested voters. This complicated things since such voters cannot be assumed to judge the options in the light of the independent standard of the common good, but rather in accordance with the subjective standard of what is good for them. However, I argued that if all voters are assumed to be equally minimally competent when judging the

---

<sup>10</sup> See 5.2.1 above.

<sup>11</sup> See 5.2.2 above.

options in accordance with their subjective standard (equally minimally end-competent), they are on average better than chance in correctly judging the options in accordance with the independent standard of the common good. Under this latter condition, the Heterogeneous Competence CJT can be employed to give us the asymptotic conclusion that the majority is nearly certain to vote for the correct (common-interest) option. Hence the following assumption can serve in a second argument from weak collective optimality, given equal-stakes cases with sufficient numbers of self-interested voters.<sup>12</sup>

***Equal Minimal End-Competence:*** Every voter is equally better than chance at judging the options correctly in the light of her desired end, such that every voter's individual end-competence is  $e > \frac{1}{2}$ .

Starting from an equal-stakes case with self-interested voters, I then showed that by successively replacing these voters with common-interested voters the asymptotic conclusion can be preserved. Thus the case for the weighted majority rule's weak collective optimality can be made even for a mix of common- and self-interested voters.<sup>13</sup>

In an attempt to relax Equal Minimal End-Competence further, I relied on yet another Condorcet theorem. The Distribution-Neutral CJT states the asymptotic conclusion for decisions with two options and  $n$  voters, if the voters have an average competence (in the light of the common good) of at least  $\frac{1}{2} + \frac{1}{2n}$ , as well as vote independently and according to their judgment. I argued that a voter's competence in the light of the common good (her CJT-competence) equals her end-competence if she is a majority stake-holder or a common-interested minority stake-holder, and equals her end-*in*competence if she is a self-interested minority stake-holder. Thus, it suffices to assume that the voters have *any* end-competences such that the following holds.

***Raised Average CJT-Competence:*** The voters' average CJT-competence  $c^*$  exceeds chance by at least  $\frac{1}{2n}$ , that is,  $c^* \geq \frac{1}{2} + \frac{1}{2n}$ .

Thus, for equal-stakes with self- or common-interested voters or a mix thereof, the weighted majority rule can be shown to be weakly collectively optimal on the rather weak competence assumption that their end-competences result in an average CJT-competence of at least  $\frac{1}{2} + \frac{1}{2n}$  — again, given a sufficiently large number of voters. This minimum level of average CJT-competence decreases when the number of voters increases and approaches  $\frac{1}{2}$  as number of voters  $n$  approaches infinity. Thus, the assumption is rather undemanding for very large groups of voters.

Stating the assumption in terms of the voters' CJT-competence has a further upshot: it eliminates the need for assuming self- or common-interested

---

<sup>12</sup> See 5.2.3 above.

<sup>13</sup> See 5.2.4 above.

motivation. Instead, we can now state that the voters may have any desired ends, as long as their end-competences result in an average CJT-competence level of at least  $\frac{1}{2} + 1/2n$ .

In unequal-stakes cases, voters are assigned unequal numbers of votes by the weighted majority rule. Thus, in these cases, the rule is not extensionally equivalent with the simple majority rule, for which the Condorcet theorems are designed. I suggested that we could redescribe these cases to fit the theorems by treating each vote as a separate voter. However, this implies that the theorems' independence assumption is violated, since votes are assigned within indivisible vote bundles. Thus, if one vote within such a bundle is cast for an option, the others must be cast in the same way. They are thus probabilistically correlated (dependent).<sup>14</sup>

In order to deal with this difficulty, I employed yet another theorem, the Voter Dependence CJT. By restating this theorem in terms of votes rather than voters, as Vote Dependence CJT, it can be employed in an argument to the effect that, given Minimal Average CJT-competence, the majority of votes is nearly certain to be cast for the common-interest option, when the number of votes is sufficiently large *and* when the votes are on average tolerably correlated. The additional last clause renders this argument from weak collective optimality still weaker than its predecessors. On the other hand, it gives us a way to handle dependence due to indivisible vote bundles, and thus to extend the argument to unequal-stakes cases.<sup>15</sup>

Unfortunately, indivisible vote bundles are not the only source of correlation between votes, which violates the independence assumption. There might also be other common causes, such as shared evidence, to the same effect. To be sure, correlation due to common causes may be handled by the above Vote Dependence CJT as well. However, for cases featuring both indivisible vote bundles and common causes, correlation may easily rise beyond tolerable levels. It has been proposed that independence between votes could be preserved by restating the independence assumption conditional on all common causes. This proposal results in another theorem: Problem-Specific CJT. I considered this theorem in some detail but was forced to conclude that it cannot be properly applied to my unequal-stakes cases. The main reason was that the theorem presupposes that there be no direct dependence between votes (but only between votes and common causes). This condition is violated in unequal-stakes cases, where there is direct dependence between votes within a vote bundle. A stronger case for the weighted majority rule is therefore pending on further development of Condorcet theorems that accommodate correlation between votes.<sup>16</sup>

---

<sup>14</sup> See 5.2.6 above.

<sup>15</sup> See 5.2.7 above.

<sup>16</sup> See 5.2.8 above.

My argumentative strategy of Chapter 5 has been to apply different Condorcet jury theorems to a series of decision-making contexts — from equal-stakes common-interested voters cases to unequal-stakes cases with mixed voter motivation. The theorems' respective asymptotic conclusions have then been employed as premises for arguments from weak collective optimality, thereby adapting these arguments to these different contexts. For all of the stated contexts, it was shown that full voter competence (in the light of the voters' respective desired ends) is not required to establish the weighted majority rule's weak collective optimality. The upshot of these arguments is thus that the above strong Competence assumption can be considerably relaxed. However, this is done at the cost of settling for weak rather than strong collective optimality and of relying on additional assumptions.

This argumentative strategy of adapting the argument from weak collective optimality to less and less constricted contexts, in a series of steps, has had the advantage of clarifying the price we have to pay for this adaptation. E.g., we can see that moving from equal-stakes to unequal-stakes cases (with mixed motivation) introduces the need of an additional assumption, requiring a tolerable average correlation between votes, to retain the weighted majority rule's weak collective optimality.

An additional upshot of my application of different Condorcet jury theorems within the present study was that these theorems are less demanding than commonly thought. Usually, they are taken to rest on the assumption that the voters share a common goal (e.g. they are all truth-trackers). My arguments showed that this assumption can be relaxed: voters need not be assumed to share any common goal (such as finding out the truth — or indeed, identifying the common-interest option) but may have their own self-interest (or, indeed, any interest or cause whatsoever) as their goal. What is required is solely that their competence regarding this goal translates into the required level of (average) CJT-competence. In addition, my arguments showed that these Condorcet jury theorems are relevant not only for the simple majority rule, but also for the novel weighted majority rule.

When it comes to the practical relevance of my results in Chapter 5, it can be noted that all the therein stated arguments for the weighted majority rule rest on the condition that there are sufficiently many voters (or votes). This means that within the traditional political domain of elections and referenda, with large electorates, the weighted majority rule emerges as a promising candidate from a common-good perspective. Thus, my results seem relevant for e.g. the 2012 French presidential election, with millions of voters, or the Stockholm congestion tax referendum, with hundreds of thousands.<sup>17</sup> These appear to constitute appropriate contexts for close-to-optimal outcomes by weighted majority rule.

---

<sup>17</sup> Cf. 1.1 above.

Whether my arguments are relevant in these cases depends of course also on whether the relevant competence and correlation assumptions are satisfied by the real-life voters. I do not have empirical data to settle this latter question for the stated two cases and thus leave this question for further research.

### 8.3.3 The better-than-the-average-voter argument

All my arguments from weak collective optimality were stated on the assumption that there are sufficiently many voters. Yet when smaller groups vote — say, the customers of the local pub planning a special sports night<sup>18</sup> — there might not be enough voters. Then, weak collective optimality would still be preserved for quite high CJT-competence levels. However, in Chapter 6 I also proposed an alternative argument for such contexts, by suggesting how the non-asymptotic conclusions of the considered Condorcet theorems could be employed in an argument for the weighted majority rule that does not rely on large numbers of voters. The *better-than-the-average-voter argument* asserted that this rule is better than the average voter's vote at selecting the common-interest option when the average levels of correlation between votes are below a given threshold.<sup>19</sup>

Still, outperforming the average or random voter's vote appears not to be an overly recommending feature for a democratic decision rule, as this seems far from generating collectively optimal outcomes, on any plausible interpretation of 'collectively optimal'. Moreover, relying on a random vote is rarely a practically relevant alternative when making a collective decision. In reality, if we would set out to rely on any single vote, we would most likely opt for the vote or judgment of an expert (who has good knowledge of the overall effects of the options). And such an expert's competence might easily outperform the average voter's.

On the other hand, in real-life decisions (such as the pub vote), we may not be able to identify such an expert. Alternatively, it might be the case that all voters are about equally CJT-competent. Then the weighted majority rule may be a valuable tool of decision-making even for smaller groups, as suggested by the *better-than-the-average-voter argument*.

### 8.3.4 The behavioural argument from weak collective optimality

In Chapter 6 I moreover used some of Chapter 5's results to state an argument from weak collective optimality from an assumption concerning voting behaviour, rather than concerning the motivational set-up of the voters. The *behavioural argument* shows that the majority of votes is nearly certain to be

---

<sup>18</sup> Cf. 1.1 above.

<sup>19</sup> See 6.3 above.

cast for the common-interest option, given that the number of votes is sufficiently large and that the votes are on average tolerably correlated and given the following assumption.

**Minimal Average Reliability:** The votes are minimally reliable on average, that is, the average of their individual probabilities to be cast for the correct option (according to the independent standard) is above chance,  $r^* > 1/2$ .

Stating the assumption in terms of voting behaviour, rather than in terms of the desires and beliefs (or judgments) of the voters, allowed us to make a case for the weighted majority rule without assuming the voters to display any specific mental structure.<sup>20</sup>

Note that the *behavioural argument* is closely related to the previous *extended argument from collective optimality*, in the following way. With its Minimal Average Reliability assumption, the behavioural argument features a relaxed version of the latter argument's Self- Or Common-Interested Voting assumption. Instead of requiring all voters to vote (reliably) according to their self-interest or the common interest, Minimal Average Reliability allows voting in accordance with *any* end, under the constraint that this results in votes being on average more likely than not to be cast for what is in fact the correct (or common-interest) option. We can now see that the price we pay for this weaker assumption is that the result of the behavioural argument is stated in terms of weak collective optimality and conditional on tolerable average levels of correlation (where the extended argument showed the weighted majority rule to be strongly collectively optimal without this further condition).

## 8.4 The scope of the argument can be extended to multi-option decisions

In Chapter 7 I extended the scope of the case for the weighted majority rule from binary decision-making to multi-option contexts. This required rendering the weighted majority rule in a version that can handle more than two options at once. I proposed the following definition.

**The Condorcet-Style Weighted Majority Rule:** For all individuals and any decision, for each pair of options, (a) every individual is assigned a number of votes in proportion to her stakes for this pair, (b) an option  $x$  is collectively ranked above an option  $y$  if and only if  $x$  receives more votes than  $y$  in a pairwise vote, and (c) the top-ranked option is selected as the outcome.

---

<sup>20</sup> See 6.2 above.

I then argued that this extended rule is collectively optimal even in multi-option contexts, given Self- Or Common-Interested Voting, according to the following.

#### 8.4.1 The further-extended argument from collective optimality

Recall that the above *generic* and *extended* arguments from collective optimality show that the weighted majority rule selects the collectively better option in any binary decision (given certain assumptions). We could also say that this rule ranks the collectively better of the two options above the other. Since this holds for every binary decision, and since the Condorcet-style weighted majority rule, in a way of speaking, assembles a collective ranking from all possible binary rankings of the given options, we can infer that the latter rule ranks all the options according to how collectively good they are, and hence selects the (top-ranked) collectively best option. I stated that the conclusion of this *further-extended argument from collective optimality* is that the weighted majority rule is collectively optimal even in multi-option settings (given the same set of assumptions).<sup>21</sup>

The *further-extended argument* presupposes Self- Or Common-Interested Voting. I pointed out that, in multi-option decisions, this assumption is ambiguous. It can be interpreted as either one of the following two claims.

***Local-Scope SCI-Voting:*** Every voter votes for the option that is in her self-interest or in the common interest, *within each binary decision*. That is, every voter votes according to her *local* self-interest or to the *local* common interest.

***Global-Scope SCI-Voting:*** Every voter votes such as to promote the option that is in her self-interest or in the common interest, *among all the available options*. That is, every voter votes according to her *global* self-interest or to the *global* common interest.

I argued that the *further-extended argument* presupposes the Local-Scope interpretation. Yet since any plausible interpretation of individual self-interest concerns how the individual fares on the whole (a whole life or at least a considerable stretch of time), it seems that Local-Scope SCI-Voting implies that a voter votes against her self-interest, under certain circumstances. This means that the case for the weighted majority rule in multi-option settings derives the winning common-interest option from input information about either what is in — or what is against — the voters' self-interest, depending on whichever makes the rule perform better. Obviously, this renders the case for the weighted majority rule much less promising.<sup>22</sup>

---

<sup>21</sup> See 7.2 above.

<sup>22</sup> See 7.3 above.

For this reason, I dedicated the main part of Chapter 7 to exploring the implications of adopting Global-Scope SCI-Voting instead. This assumption can, again, be understood as implied by the conjunction of the Self- Or Common-Interested Voting, Competence and Success assumptions (if Competence is understood as requiring correct beliefs about which way of voting serves one's end, rather than about which option does).

First, I considered exclusively self-interested voters. I showed that under Global-Scope SCI-Voting, logrolling becomes possible. Logrolling is strategic voting among several voters who 'trade' votes with each other on different pairwise decisions (thus occasionally voting against their self-interest in the local-scope context of such a binary decision). I showed that logrolling might not be stable since the logrollers might face what I called the logrollers' dilemma, where defecting from logrolling is better for each (although all logrollers would be better off if no one defected). This was shown with several *backward induction arguments*, on a rather strong assumption of common knowledge.

**Common Knowledge:** Every voter knows, and knows that everyone knows, and so on, (a) about the upper limit — or end — of the sequence of decisions they face and (b) about every voter's satisfying Self- Or Common-Interested Voting — by being self-interested — as well as Competence and Success.

Still, if this assumption is relaxed and certain levels of uncertainty (regarding the upper limit or end of the sequence of decisions) are allowed, logrolling may nevertheless be stable — and undermine the collective optimality of the weighted majority rule.<sup>23</sup>

I then considered a proposal on how to avoid logrolling, by including all options on which votes might be traded in one single decision. Such a decision would offer compound options that comprise all possible combinations of simple options on which vote-trading might occur. I show that this move does not succeed and thus cannot save the weighted majority rule from the damaging effects of logrolling.<sup>24</sup> Moreover, for compound four-option decisions (derived from two binary decisions on which vote-trading might occur), the logrollers are shown to escape the logrollers' dilemma, even in the absence of Common Knowledge. This thus speaks against devising compound options.<sup>25</sup> However, I also argue that, given compound options, logrolling becomes somewhat less profitable for logrollers and somewhat less damaging regarding collective optimality. This is so since the logrollers can at best achieve a cycle or tie, that is (given some random tie- or cycle-breaker) a gamble on sets of options that include options which make them

---

<sup>23</sup> See 7.3.1 above.

<sup>24</sup> See 7.3.2 above.

<sup>25</sup> See 7.3.5 above.

worse off than the outcome from non-strategic voting.<sup>26</sup> For compound four-option decisions, I show that this means that logrolling is less profitable for the logrollers and less collectively damaging than logrolling in the separate binary decisions the compound case is made up of. This speaks in favour of devising compound decisions.<sup>27</sup>

However, there is another worry with such a move: it results in an exponential increase in pairwise decisions between these compound options (compared to the number of pairwise decisions between all the simple options from which the compound ones are derived). This would require incessant voting which — as voting is costly — would increase the procedural costs of decision-making. Moreover, if less than full competence is assumed, devising compound (e.g. conjunctive) options may decrease voter competence and hence generate much lower group competence (compared to retaining simple options), as was shown previously.<sup>28</sup> Finally, as the number of options increases, the probability of a correct collective ranking has been shown to decrease. This is so since this probability is the product of the probabilities of a correct binary ranking for each pair of options. When the latter probabilities are below one, the probability of a correct collective ranking of all the options decreases for every option we add.<sup>29</sup>

Considering all these results, devising compound options appears not to be a good solution to avoid logrolling. This result should advise the local pub staff not to let its customers vote on compound options such as *football&entrance-fee&smoking* vs. *golf&entrance-fee&smoking*, etc.<sup>30</sup>

Besides logrolling, I have shown that there is another way of strategic voting. This is practiced by individual voters who in a series of binary decisions may distribute or 'push' their votes between the options such as to achieve an outcome that is better for them than the outcome from non-strategic voting. Since single voters have no incentive to defect from such a chosen strategy, they cannot find themselves in situations like the logrollers' dilemma. Hence, individual strategic voting is stable and constitutes a problem for the weighted majority rule's collective optimality, even in the absence of logrolling opportunities.<sup>31</sup> I have, however, also shown that any single voter can at best achieve a cycle, that is, a gamble on sets of options that include options which makes her worse off than the outcome from non-strategic voting.<sup>32</sup> Still, the upshot of my arguments so far is that, given Global-Scope SCI-Voting, and more specifically, assuming voters to be self-

---

<sup>26</sup> See 7.3.3 above.

<sup>27</sup> See 7.3.4 above.

<sup>28</sup> See 6.4 above.

<sup>29</sup> See 7.7 above.

<sup>30</sup> Cf. 1.1 above.

<sup>31</sup> See 7.4 above.

<sup>32</sup> See 7.4.1 above.

interested, strategic voting is possible and may undermine the collective optimality of the weighted majority rule.

I also showed that this conclusion does not hold for common-interested voters who satisfy a weaker competence assumption,

**Weaker Common Knowledge:** Every voter knows, and knows that everyone knows, and so on, that every voter satisfies Self- Or Common-Interested Voting — by being common-interested — as well as Competence and Success.

However, if we introduce some level of uncertainty regarding the other voters' ends and competence, strategic voting resurfaces again and spells trouble for collective optimality of the weighted majority rule.

In the face of these problems from strategic voting, I pointed out the need to investigate whether the Condorcet-style weighted majority rule could be rendered in an alternative version, or equipped with an alternative tie- and cycle-breaker, in order to make it (more) strategy-proof.<sup>33</sup> However, in the absence of such improvements, it must be concluded that the case for the weighted majority rule in multi-option settings is conditional on a myopic kind of self- or common-interested voting-behaviour that occasionally implies voting against the voters self- or common-interest, as seen from a larger picture. In real-life decision-making, this is hardly a plausible assumption. E.g. in the local pub decisions on whether to televise the golf tournament or the football match, whether to raise an entrance fee and whether to allow smoking, it seems futile to expect that voters would not see that they could trade their votes on different issues with each other, but vote appropriately (non-erratically) on every binary decision in isolation.

To conclude: the good news is that the argument for the weighted majority rule's collective optimality can be extended from binary to multi-option decisions. Yet, sadly, by requiring input information about either what is in — or what is against — the voters' self-interest, depending on whichever makes the rule perform better, this case seems somewhat less compelling, and also less practically relevant, than one could have hoped for.

## 8.5 Conclusion

Each of the arguments for the weighted majority rule proposed in this study specified a number of assumptions that are jointly sufficient for this rule's (weak or strong) collective optimality. In other words, each argument delimits a set of cases of collective decision-making in which this rule selects the collectively best option (with certainty or near certainty). By relaxing the assumptions, along my three proposed lines of inquiry, I have shown that the

---

<sup>33</sup> See 7.5 above.

case for the weighted majority rule can be made for a considerably larger set of cases than previously thought.

I have, however, not been concerned with identifying the limiting cases for which this result no longer applies. That is, I have not set out to find the necessary conditions for the weighted majority rule's collective optimality. Hence, it may turn out eventually that this result can be attained under entirely different assumptions, and thus possibly for still other sets of collective decision-making cases. This question, along with others that have been raised but not answered by this study, I must leave for another day.



# Appendix

## A.1 Introduction

In this study, I have among other things claimed that a case for the weighted majority rule can be made on an assumption that voters are self-interested. From this, several further questions might arise: would *accepting* the weighted majority rule be in accordance with the individual's self-interest? By 'accepting', I mean here 'agreeing to its implementation and to the collective enacting of its outcomes'. Moreover, even if a self-interested voter could be assumed to accept the rule, would *voting* be in accordance with her self-interest?

Douglas Rae ponders the question of which among all possible voting rules would a self-interested voter accept. His answer is: the simple majority rule. This rule promotes the well-being of each individual. This gives each self-interested individual a reason to consent to it.<sup>1</sup> At least, this holds when the voter is understood as an anonymous constitution-maker who is placed behind a veil of ignorance.

As pointed out before, simple majority rule can be understood as an instance of the weighted majority rule for cases where all stakes are equal. I show in this appendix that Rae's result can be extended to cover the weighted majority rule as well.

I state my arguments in an appendix, rather than the main text of this study, as Rae's arguments proceed from a normative criterion that differs from my own: Rae sets out to evaluate voting rules from an individual well-being perspective and employs a criterion of individual consent. The present study has a different normative outlook: it sets out to argue for a voting rule from a common-good perspective. Thus, the question of individual consent is not of any direct relevance for the main task of this study: showing under which assumptions the weighted majority rule is collectively optimal. Still, it is closely linked to this study by relying on a self-interested voter assumption. Moreover, Rae's argument lends itself to a case for the weighted majority rule that is interesting in its own right.

---

<sup>1</sup> Rae (1969). See even Taylor (1969).

In A.2, I reconstruct Rae's argument. In A.3, I connect his argument to the present study. I argue that some of Rae's assumptions can and should be altered and that the result of such alterations is an improved argument that actually advocates the weighted majority rule. In A.4, I tie my results to two well-known paradoxes of democracy: *Wollheim's paradox* and the *voter paradox*. Finally, section A.5 concludes.

## A.2 Rae's argument for an individually optimal voting rule

Rae analyses a variety of voting rules from an individual, self-interested perspective.<sup>2</sup> For his analysis, he focuses exclusively on a constitution-maker who is to choose among all possible voting rules according to which one is overall best for him.

Let us take a look at the different parts of Rae's analysis. First, the *constitution-maker* is 'a single, anonymous individual who wishes to optimize the correspondence between his own values, however selfish or altruistic, and those expressed by collective policy. This individual would like to "have his way" as often as possible, by securing the adoption of proposals he likes and the defeat of proposals he dislikes'.<sup>3</sup>

Second, the constitution-maker is claimed to choose among *all possible* voting rules. To be exact, though, Rae assumes that there are exactly  $n$  voting rules, for any group of  $n$  voters ( $n \geq 3$ ), who face a binary decision between supporting some policy  $x$  and defeating  $x$  (which means preserving status quo). At the one extreme of these  $n$  rules Rae locates the 'rule of consensus':  $x$  is passed only if all  $n$  voters support it. At the other extreme is the 'rule of individual initiative':  $x$  is passed only if one voter supports it.<sup>4</sup> In between these extremes then are the rules which state that  $x$  is passed only if  $n-1$  (or  $n-2$ ; ...; or  $n-(n-2)$ ) voters support it. Thus, Rae considers neither, e.g. non-anonymous, random or weighted voting rules nor specific rules for non-binary decision-making.

So among these  $n$  voting rules, which should the constitution-maker choose? Rae seeks to give a general answer to this question, one that is not limited to certain sets of options or groups. So he places the constitution-maker behind a *veil of ignorance*, knowing 'nothing about the (long-run) agenda which will confront the [group], about the ways individuals will evaluate the proposals which do arise, or about the factional structure of the [group]'.<sup>5</sup> What he knows is solely that each voter (including himself) will

---

<sup>2</sup> Rae (1969).

<sup>3</sup> Rae (1969: 41).

<sup>4</sup> Rae (1969: 49).

<sup>5</sup> Rae (1969: 41).

either support or reject each proposal, and that each one's support or rejection is independent of any other's.<sup>6</sup>

Behind his veil of ignorance the constitution-maker thus knows only that he will face one of four possible events for each decision:

- (A) A policy he supports is collectively defeated.
- (B) A policy he opposes is collectively passed.
- (C) A policy he opposes is collectively defeated.
- (D) A policy he supports is collectively passed.

In the first two events — (A) *supported but defeated policy* and (B) *opposed but passed policy* — the 'values' expressed by the outcomes of the decision do not correspond to the constitution-maker's own. In the last two — (C) *opposed and defeated policy* and (D) *supported and passed policy* — they do correspond.

According to the initial characterisation of the constitution-maker, he wants to 'have his way' as often as possible. This goal Rae then specifies as an individualist normative criterion: 'One should choose that decision-rule which minimizes the sum of the expected frequencies for (A) in which the committee does not impose a policy which his value schedule leads him to support, and (B) in which the committee imposes a policy which his value schedule leads him to oppose'.<sup>7</sup>

The above description generates a model within which the expected frequencies of (A) and (B) can be calculated for any voting rule that requires  $n - m$  voters to support a policy in order to pass it. Rae assumes that  $n > m \geq 0$ . This means that the group will not vote on any policy that does not have any supporters at all among its members.

Informally stated, the results are as follows. Under the *rule of consensus*, the expected frequency of (A) *supported but defeated policy* is at its maximum, as a policy is passed only if everyone supports it. The expected frequency of (B) *opposed but passed policy* is, however, zero since the constitution-maker's rejection will suffice to defeat a policy. Under the *rule of individual initiative*, in contrast, the tables are turned. The expected frequency of (A) is zero since the constitution-maker's support will suffice for a policy to pass. The expected frequency of (B) is, however, quite high, as a policy is defeated only if no one supports it. Between these two extremes, the expected frequencies for event (A) are monotonically increasing with the number of individuals whose support is required for a policy to pass, while the expected frequencies for event (B) are monotonically decreasing. The frequency curves are thus opposed.

---

<sup>6</sup> The additional initial assumption that everyone is equally likely to support or oppose any policy is subsequently relaxed in Rae's appendix (Rae 1969: 55–56).

<sup>7</sup> Rae (1969: 42).

The normative criterion requires that the *sum* of the expected frequencies of (A) and (B) is minimised for the constitution-maker.<sup>8</sup> Rae shows that the optimal decision rule is located between the two extremes *rule of consensus* and *rule of individual initiative*. For an odd number of group members, the sum is minimised when the required number of supporters to get a policy passed is  $(n+1)/2$ . For an even number, this minimum occurs both at  $n/2$  and at  $n/2+1$ .<sup>9</sup>

*Simple majority rule*, requiring that more than half the voters support a policy in order for it to pass, is one optimal voting rule according to Rae's result (although it would not pass a policy if exactly  $n/2$  voters support it). Thus, Rae concludes that 'majority-rule is as good (i.e. optimal) as any alternative decision-rule, given the model proposed here'.<sup>10</sup> However, we should note that a voting rule requiring that at least half the voters support a policy in order for it to pass is equally good. We may call this the *non-minority rule*.

What Rae's argument shows is thus that any one of these two voting rules may be chosen by the constitution-maker under the described circumstances. For a given decision, they maximise the probability that the constitution-maker will 'have his way'. In the long run, this means that they will ensure the constitution-maker's preferred outcomes as often as possible.<sup>11</sup>

### A.3 How Rae's argument is relevant in the present context

As stated above, Rae's constitution-maker is 'a single, anonymous individual who wishes to optimize the correspondence between his own values, however selfish or altruistic, and those expressed by collective policy'.<sup>12</sup> In the terminology of the present study this means that he is either self-interested or common-interested — when it comes to first-order collective decision-making. Yet when it comes to the second-order choice of a collective decision rule, he is purely self-interested. He seeks maximum correspondence between his first-order preferences (whatever they may be) and the collective outcome. This is his normative criterion.

Thus, Rae's argument provides an answer to the question of which voting rule a self-interested individual would accept. His answer is simple majority rule (or, we may want to add, non-minority rule). This does not square well

---

<sup>8</sup> The expression 'for the constitution-maker' is added since the events (A) and (B) make an indexical reference to 'his value schedule' (Rae 1969: 42, cf. also above).

<sup>9</sup> See Rae (1969: 47–51), who also illustrates this with a number of tables and graphs.

<sup>10</sup> Rae (1969: 52).

<sup>11</sup> Rae (1969: 42). For a general mathematical proof of this argument, see Taylor (1969).

<sup>12</sup> Rae (1969: 41).

with our hope of finding a prudential justification of the weighted majority rule. However, as we will see shortly, Rae's argument rests on a number of controversial assumptions. Relaxing some of these assumptions brings us much closer to our goal of giving a prudential justification of the weighted majority rule.

For starters, there is the problem that Rae's argument assumes that for every policy  $x$  and every voter  $i$ ,  $i$  will either support or reject  $x$ . Thus, there is no room for indifference. A voter would be indifferent if he would rank  $x$  as just as good as the status quo. This possibility should surely be granted (even on the collective level). Second, framing the argument in terms of binary decisions between supporting and rejecting policies makes all decisions status-quo dependent. Hence, the argument does not apply to decisions between two options of which neither is the status quo. So it cannot account for cases where several mutually exclusive policies are collectively ranked above status quo and only one of them has to be selected. Third, the argument cannot make sense of 'the problem of intensity', the idea that some decisions might be more important for some voters than for others. In our terminology: in some decisions there might be more at stake than in others. Again, the argument should be able to account for that. These problems are addressed in the following section. There we start by considering a further problem with Rae's assumptions, one which he explicitly acknowledges: the problem of bias.

### A.3.1 The problem of bias

Rae acknowledges that his argument for simple majority rule is based on the assumption that the constitution-maker is not biased in favour of one of the options. Allowing bias significantly changes the conclusion of the argument. Consider, e.g. a voter with a general conservative bias in favour of the status quo.<sup>13</sup> This means that he would assign more disvalue to (B), i.e. 'bad actions', than to (A), i.e. 'bad inactions'.<sup>14</sup> But then his normative criterion would no longer be to minimise the sum of the expected frequencies of (A) and (B). Rather, the criterion would have to be adjusted: it would then advocate as optimal a decision rule which minimises the sum of the *weighted* expected frequencies of (A) and (B), with the weights chosen in proportion to the assigned (dis)values. Rae shows that if weights of 1 and 2 are chosen respectively for (A) and (B), then the optimal voting rule would be what we may call a *two-thirds majority rule*: a policy is passed only if 2/3 of the  $n$  voters support it.

---

<sup>13</sup> Rae characterises this bias as a '*positional* (as opposed to substantive) *preference*' (1969:63).

<sup>14</sup> Rae (1969:63).

What do the weights of 1 and 2 signify? Rae repeatedly stresses ‘the enormous difficulty of supplying meaningful quantities for [these weights]’, but concedes that, ‘[o]n the assumption that the weights themselves make sense’, we arrive at the adjusted normative criterion, and through that at some alternative optimal voting rule.<sup>15</sup> There are two separate problems that Rae hints at: one is the problem of whether the weights can be rendered meaningful, in the sense of deriving them from a meaningful and relevant difference between the two events ‘bad action’ and ‘bad inaction’. This difference is simply how good (bad) they are, in comparison, for the constitution-maker. The other is the problem of correctly assessing this difference in order to assign the proper weights. Both problems have to be solved in order to assign non-arbitrary weights and by means of them derive a non-arbitrary, optimal voting rule.

From this description, we can see that the problem of bias in effect breaks down to an instance of what Rae calls the ‘problem of intensity’: that there may be varying stakes in different decisions. Rae puts the problem aside as theoretically intractable, although important.<sup>16</sup> The present study does not share Rae's pessimism. To say the least, *if* we can make sense of e.g. the idea that bad action (B) is twice as bad as bad inaction (A), we can conclude that the weights of 2 and 1 should be assigned respectively. Let us see whether within the present framework, Rae's argument can be advanced some steps further, if varying stakes are allowed.

Note first that an additional problem with Rae's framework is that the now proposed adjusted normative criterion only holds for a constitution-maker who is always and only biased in a certain way — e.g. conservatively. But it may also be the case that he is only sometimes conservatively biased — for instance when it comes to family issues. In other domains, e.g. concerning education and the sciences, he may instead have an anti-conservative (innovative) bias. Since, as assumed, he does not know which issues he will face, and thus how his biases will hit, he is at a loss to evaluate different decision rules, even with the adjusted normative criterion.

Alternatively, this objection can be framed in terms of the veil of ignorance. In Rae's framework, the veil is here partially lifted and allows a glimpse on the constitution-maker's general bias. However, if we want to preserve the generality of the constitution-maker's choice, we should make any possible bias invisible behind the veil. The challenge is now to allow the constitution-maker to find one single optimal voting rule for all decisions, given that he does not know his own bias. More generally speaking, we want to assume that he does not know what stakes he will have in these decisions.

Allowing all kinds of possible biases — even those not defined in terms of the status quo — and more generally speaking, allowing varying stakes,

---

<sup>15</sup> Rae (1969: 53).

<sup>16</sup> Rae (1969: 41, footnote 6). Cf. even Rae and Taylor (1969).

means that we do not have to frame the outcomes or events that the constitution-maker faces in terms of *passed* or *defeated* policy  $x$  (both having the status quo non- $x$  as their base line). A more general, status-quo independent way of framing the outcomes is in terms of the collective ranking of any two options  $x$  and  $y$ . Then, either this ranking corresponds to the constitution-maker's individual ranking of the options in question, or it does not.

This has the further advantage of reducing the number of events for the constitution-maker, facing options  $x$  and  $y$ .

**(I) Correspondence:** the constitution-maker ranks  $x$  above  $y$ , as does the collective ranking.

**(II) Non-correspondence:** the constitution-maker ranks  $x$  above  $y$ , but the collective ranking does not.

(I) is equivalent to the union of (C) *opposed and defeated policy* (if we consider  $y$ ) and (D) *supported and passed policy* (if we consider  $x$ ). (II) is equivalent to the union of (A) *supported but defeated policy* (if we consider  $x$ ), (B) *opposed but passed policy* (if we consider  $y$ ) and an additional event (E). (E) covers all cases of non-correspondence in which the constitution-maker is not indifferent between  $x$  and  $y$  (that is, supports or opposes some of them), while the *collective ranking is indifferent* (that is, forms a tie). This possibility was missing in Rae's model.

With this representation of the problem, we can also account for the other missing event:

**(III) Individual indifference:** the constitution-maker is indifferent between  $x$  and  $y$ , while the collective ranking either ranks one above the other or is indifferent as well.

### A.3.2 Deriving the weighted majority rule from the Raean framework

What would be the proper normative criterion for a self-interested constitution-maker facing these three events? His main objective would be to have as much correspondence and as little non-correspondence as possible. Indifference events do not matter to him — at least as long as his goal of ‘having his way’ is not a fetish (such that he seeks collective indifference whenever he is indifferent), but rather the goal to get what he prefers when this matters to him (that is, when he is not indifferent). However, as here assumed, not all correspondence will be equally good and not all non-correspondence equally bad for the constitution-maker. Some may be much better, and some much worse, due to bias — or simply due to the stakes involved in different decisions. So his objective would be something like ‘having as much correspondence as possible — and especially in high-stakes decisions — and as

little non-correspondence as possible — again, especially in high-stakes decisions’.

One way of clarifying this clumsily stated idea is to say that the constitution-maker wants to maximise the difference between the weighted expected frequency of (I) *Correspondence* minus the weighted expected frequency of (II) *Non-correspondence*, with weights chosen in proportion to whatever the stakes involved. This then is the improved, *stake-sensitive normative criterion*. It is improved because it is more general, applying even to a constitution-maker behind a bias- and stake-proof veil of ignorance.

(We can see from the criterion's formulation that we might as well include event (III) *Individual indifference* in it. Since this event occurs only when the constitution-maker is indifferent, his stakes will be zero, and hence the weights will be zero. Event (III) is thus easily seen to be irrelevant within this extended version of the stake-sensitive normative criterion.)

Now this seems to be a hopeless criterion for finding one single optimal voting rule for all possible decisions, since the stakes — and thus weights — may vary for every instance of events (I) and (II). But it helps if the constitution-maker shifts perspective. He knows that he will face an indefinite number of decisions. And he knows that he will have some number of stakes in each, ranging between zero (indifference) and the total amount of stakes in the decision,  $s$  (i.e. the sum-total of all the voters' well-being differentials — thus, holding  $s$  stakes, he would be the only one affected by the decision). That is, he knows that for every decision, his stakes can be any number of these  $s$  stakes. What he wants can now be re-described as having as many of his stakes in instances of event (I) and as few in instances of event (II) as possible. Thus, for every one of his stakes, he wants to maximise the probability that it ‘occurs’ in (I), and — equivalently — minimise the probability that it ‘occurs’ in (II). So in a slight reformulation, his normative criterion is to minimise the expected frequency of (II) for each of his stakes (rather than for himself). This we may call the *stake-centred normative criterion*.

(That a voter  $i$ 's stake ‘occurs’ in (I) [or ‘occurs’ in (II)] simply means that it is one of  $i$ 's stakes in a decision where the option with higher [lower] well-being for  $i$  is selected.)

I now want to assimilate this criterion to Rae's original criterion to facilitate the assimilation of his proof to my purposes. As stated before, (II) *Non-correspondence* is equivalent to the union of (A) *supported but defeated policy*, (B) *opposed but passed policy* and (E) *supported or opposed but collectively indifferent policy*. Rae's criterion only considers (A) and (B). Yet by focusing only on decisions with an odd number of stakes, such that there cannot be any collective indifference, (E) can be expelled. So for odd-stakes cases, (II) is equivalent to the union of (A) *supported but defeated policy* and (B) *opposed but passed policy*. Rae's normative criterion requires that the sum of the expected frequencies of (A) and (B) is minimised for the constitution-maker. Thus, it is equivalent to my stake-centred normative

criterion, except that it focuses on the constitution-maker — as one of  $n$  voters — instead of on each of his stakes — as one among a total of  $s$  stakes.

Rae shows that the optimal decision rule, according to his criterion, requires that there are  $(n+1)/2$  voters in favour of  $x$  to get  $x$  passed, for an odd number of group members. Along the same lines, we can now conclude that the optimal decision rule according to the stake-centred normative criterion requires that there are  $(n+1)/2$  stakes ‘in favour’ of  $x$  to get  $x$  passed, for an odd number of stakes. That is, a simple majority of stakes ‘in favour’ is required for a policy to be selected. And this is ensured by the weighted majority rule, which assigns numbers of votes to voters in proportion to their stakes in the decision and selects the option that gets a simple majority of votes as winner.

(That a stake is ‘in favour’ of [or ‘against’] an option  $x$  means that it is one of  $i$ 's stakes in a decision where  $x$  is better [worse] for  $i$ .)

What about decisions with an even number of stakes? There are three different possibilities here. (1) There is no collective indifference, that is, there is a majority of stakes in favour of one of the options. Then, the Raean assimilated argument holds even here. (2) There is collective indifference between two options — and the constitution-maker is indifferent as well. Then, he does not care which voting rule is employed. The weighted majority rule (with some tie-breaker) is as good for him as any other. (3) There is collective indifference between two options  $x$  and  $y$  — and the constitution-maker is *not* indifferent. This is event (E), *supported or opposed but collectively indifferent policy*. Of course, behind his veil of ignorance he does not know which of the options his stakes will be in favour of. Since there are as many stakes in total for  $x$  as there are for  $y$ , it is equally likely that his stakes will be in favour of either option. Hence, the weighted majority rule (with some tie-breaker), which chooses either one of the options, is as good for him as any other rule. So the weighted majority rule is better for the constitution-maker than any other rule in odd-stake cases as well as even-stake cases without collective indifference. And it is as good for him as any other rule in even-stake cases with collective indifference. He will thus choose it behind his veil of ignorance.

What has been shown is that the weighted majority rule would be chosen by a self-interested constitution-maker behind a veil of ignorance. To be clear about the underlying assumptions: the veil of ignorance is assumed to conceal which decisions the constitution-maker will face (the agenda), who among the  $n$  voters he will be, which among the  $s$  stakes he will hold and whether he will belong to some permanent minority.<sup>17</sup> Moreover, it has been implicitly assumed that the constitution-maker is risk-neutral, and that while he does consider the costs which arise from the outcomes of collective deci-

---

<sup>17</sup> For a discussion of the latter assumption, see Rae (1969: 53).

sion-making, he disregards the costs from the act of decision-making.<sup>18</sup> Under these assumptions, he will choose the weighted majority rule as optimal for him.

My argument also shows that Rae's argument can be taken a good step further. By dealing with the 'problem of intensities' and allowing varying stakes we arrive at a more general decision rule, the weighted majority rule, as the individually optimal rule for collective decision-making.

## A.4 Two paradoxes of democracy

**Wollheim's paradox.** The above argument shows that from behind a veil of ignorance it is in any voter  $i$ 's self-interest to accept the weighted majority rule. That is, according to how I defined 'accept' above, it is in  $i$ 's self-interest to agree to the implementation of the weighted majority rule and the collective enacting of its outcomes — even though such an outcome may at times be worse for her than its alternative. Thus, in a given decision between options  $x$  and  $y$ , it may be the case that  $x$  is selected by the weighted majority rule, while it is worse for  $i$  than  $y$ . One could then be tempted to say that  $x$  is in  $i$ 's self-interest and is not in  $i$ 's self-interest. But there is no contradiction involved since what is really meant is that  $x$  is in  $i$ 's self-interest from behind a veil of ignorance, while  $x$  is not in  $i$ 's self-interest from her 'unveiled' perspective.

This is then a solution to a prudential version of Richard Wollheim's 'paradox of democracy'.<sup>19</sup> The original paradox is framed in terms of ought-judgments, stating that 'if a man expresses a choice for A and the [democratic voting rule] expresses a [collective] choice for B, then the man, if he is to be a sound democrat, seems to be committed to the belief that A ought to be the case and to the belief that B ought to be the case'.<sup>20</sup>

Analogously, the present prudential version states that if a voter expresses a preference from self-interest for  $x$  and the weighted majority rule selects  $y$ , then the voter, if she believes the above Raean argument, seems to be committed to the belief that  $x$  is better for her and to the belief that  $y$  is better for her. Wollheim's own solution is to make explicit that the two inconsistent beliefs derive from two different principles (one concerning the democratic procedure and one concerning some property of the options) and thus give them different meaning.<sup>21</sup> Likewise in the present case, the voters' beliefs can

---

<sup>18</sup> See Rae (1969: 43).

<sup>19</sup> Wollheim (1962).

<sup>20</sup> Wollheim (1962: 78–79). For a summary of alternative solutions, see Graham (1976). Cf. even Estlund (1989) and Harrison (2005).

<sup>21</sup> For a problem with this solution, if the 'ought' in the last Wollheim quote is to be interpreted as 'morally ought', and a proposed remedy, see Graham (1976: 234, 236–237).

be spelled out as a belief that  $x$  is better for her from an ‘unveiled’ perspective and a belief that  $y$  is better for her from behind the veil of ignorance. There is no inconsistency in this pair of beliefs. By explaining how a voter could come to hold the latter belief, the Raean argument distinguishes its difference in meaning from the former and thus helps resolve the paradox.

However, there is still the problem that for a given decision in which  $y$  is selected by weighted majority rule yet is against some voter  $i$ 's self-interest, it is in  $i$ 's self-interest concerning this decision not to agree to the collective enacting of  $x$ , and thus not to accept the weighted majority rule. Does this not in effect mean that the prudential justification of the weighted majority rule and its outcomes from behind the veil of ignorance is rendered irrelevant whenever the rule's outcomes conflict with the voter's self-interest? And so, even if the paradox is resolved, does the *real* problem not remain?

It may now be suggested that this problem is due to the fact that the voter is assumed to act in accordance with a Local-Scope version of the motivational assumption:

***Local-Scope Self-Interested Motivation:*** Every individual desires to accept a voting rule, vote rather than abstain, and vote among the options, when this is in accordance with her perceived self-interest *in every single decision*.

However, the suggestion goes, as this assumption is less plausible than its Global-Scope counterpart when it comes to voting among the options, it is so also when it comes to accepting a voting rule (or voting rather than abstaining). The gist of this suggestion is then that under a Global-Scope version of this assumption, the problem disappears.

***Global-Scope Self-Interested Motivation:*** Every individual desires to accept a voting rule, vote rather than abstain, and vote among the options, when this is in accordance with her perceived self-interest *in the decision-context on the whole*.

Sadly, the problem does not disappear. The reason is that assuming Global-Scope Self-Interested Motivation does not commit the voter to considering the ‘decision context on the whole’ from the veiled perspective of the constitution-maker. For an unveiled self-interested voter, intending to act according to her self-interest in the decision context on the whole means that she accepts the weighted majority rule whenever it selects the for her better option — and does *not* accept it whenever this is not the case and non-acceptance leads to a better outcome.

However, as Global-Scope Self-Interested Motivation is a general assumption, the same holds for all voters. Hence, on the stated implication, the weighted majority rule will not be accepted by all, whenever there is no unanimity on a decision and non-acceptance leads to a better outcome for the

non-accepters. It seems then that the weighted majority rule quite likely will be abandoned. A prisoners' dilemma surfaces.

There is no solution to the prisoners' dilemma from within the 'unveiled' perspective — but there is one from behind the veil of ignorance. In the light of the discussed problem, clearly what is best for the constitution-maker from this perspective is that weighted majority rule be implemented *and properly enforced*. That is, that the incentive structure is changed as to not make non-acceptance better than acceptance for outvoted voters.

***The voter paradox.*** These arguments also pertain to the problem that, for a given decision, it may be in a voter *i*'s self-interest not to vote at all. This is traditionally considered to be the case when the voter's expected costs from voting are greater than her expected benefits.<sup>22</sup> (It is also the case when the voter's costs from voting are greater than the well-being differential between the two options — that is, when either the stakes are relatively low or the costs are relatively high.) The expected benefits can be calculated from the chance of being pivotal times the well-being differential for the voter between the given options. In particular, for decisions with millions of voters, chances of being pivotal are miniscule. Hence, expected benefits are easily outweighed by expected costs (which involve costs for the actual act of voting, but maybe even of gathering information about the options, etc.).

A similar argument as the above can be made here, making explicit the need for the *enforcement of voting* (rather than abstaining) from a self-interested perspective.

## A.5 Conclusions

In this appendix, I have argued that Rae's prudential justification of simple majority rule actually lends itself to support the weighted majority rule — if some of his assumptions are relaxed. The result is an improved argument, with a rather different conclusion, establishing the individual optimality of the weighted majority rule. I have then suggested how this argument could be employed to deal with two well-known paradoxes of democracy: *Wollheim's paradox* and the *voter paradox*.

---

<sup>22</sup> See e.g. Downs (1957: 267), Sen (1990: 34-36).

# Bibliography

- E. Anderson (2009) “Democracy: Instrumental vs. Non-Instrumental Value”, in T. Christiano and J. P. Christman (eds.) *Contemporary Debates in Political Philosophy*, Hoboken, NJ: Wiley-Blackwell.
- R. J. Arneson (2000) “Luck Egalitarianism and Prioritarianism”, *Ethics*, 110: 2, 339–349.
- R. J. Arneson (2009) “The Supposed Right to a Democratic Say”, in T. Christiano and J. P. Christman (eds.) *Contemporary Debates in Political Philosophy*, Hoboken, NJ: Wiley-Blackwell.
- G. Arrhenius (2006) “Desert as Fit: An Axiomatic Analysis”, in R. Feldman, K. McDaniel, J. R. Raibley and M. J. Zimmerman (eds.) *The Good, the Right, Life and Death: Essays in Honor of Fred Feldman*, Ashgate.
- G. Arrhenius (2010) “Democracy, Knowledge, and the Common Good”, presentation at the Tampere Club's Annual Meeting, Tampere, August 2010.
- G. Arrhenius (2011) “Defining Democratic Decision Making”, in F. Svensson and R. Sliwinski (eds.) *Neither/Nor – Philosophical Essays Dedicated to Erik Carlsson on the Occasion of His Fiftieth Birthday*, Uppsala Philosophical Studies, 58, Uppsala: UPS.
- G. Arrhenius (2013) “Democracy for the 21st Century: Research Challenges”, in Y. Elkana, S. Randeria and B. Wittrock (eds.), *Sociology Looks at the Twenty-first Century: From Local Universalism to Global Contextualism*, Brill, Leiden and Boston.
- K. J. Arrow (1963) *Social Choice and Individual Values*, 2<sup>nd</sup> edition, New Haven, CT: Yale University Press.
- D. Austen-Smith and J. S. Banks (1996) “Information Aggregation, Rationality, and the Condorcet Jury Theorem”, *The American Political Science Review*, 90: 1, 34-45.
- J. F. Banzhaf (1965) “Weighted Voting Doesn't Work: A Mathematical Analysis”, *Rutgers Law Review*, 19: 2, 317–343.
- S. Barberà (2011) “Strategyproof Social Choice”, in K. J. Arrow, A. Sen and K. Suzumura (eds.) *Handbook of Social Choice and Welfare*, volume II, Elsevier.

- S. Barberà and M. Jackson (2006) “On the Weights of Nations: Assigning Voting Power to Heterogeneous Voters”, *Journal of Political Economy*, 114, 317–39.
- B. Barry (1991) *Democracy and Power*, Oxford: Clarendon Press.
- C. Beisbart and L. Bovens (2007) “Welfarist Evaluations of Decision Rules for Boards of Representatives”, *Social Choice and Welfare*, 29, 581–608.
- J. Bentham (2000) *An Introduction to the Principles of Morals and Legislation*, Kitchener, Ontario: Batoche Books.
- D. Berend and J. Paroush (1998) “When is Condorcet’s Jury Theorem Valid?”, *Social Choice and Welfare*, 15, 481–488.
- D. Berend and L. Sapir (2007) “Monotonicity in Condorcet’s Jury Theorem with Dependent Voters”, *Social Choice and Welfare*, 28: 3, 507–528.
- K. Binmore (2007) “Interpersonal Comparison of Utility”, <http://discovery.ucl.ac.uk/14435/1/14435.pdf> (accessed 2013-08-18).
- D. Black (1958) *The Theory of Committees and Elections*, Cambridge: Cambridge University Press.
- P. J. Boland, (1989) “Majority Systems and the Condorcet Jury Theorem”, *The Statistician*, 38: 3, 181–189.
- L. Bovens (1997) “The Backward Induction Argument for the Finite Iterated Prisoner’s Dilemma and the Surprise Exam Paradox”, *Analysis*, 57: 3, 179–186.
- L. Bovens (2006) “The Doctrinal Paradox and the Mixed-Motivation Problem”, *Analysis*, 66: 1, 35-39.
- L. Bovens and W. Rabinowicz (2006) “Democratic Answers to Complex Questions – An Epistemic Perspective”, *Synthese*, 150, 131–153.
- R. B. Brandt (1998) *A Theory of the Good and the Right*, Amherst, NY: Prometheus Books.
- G. Brennan and L. Lomasky (1993) *Democracy and Decision: The Pure Theory of Electoral Preference*, Cambridge: Cambridge University Press.
- H. Brighouse and M. Fleurbaey (2010) “Democracy and Proportionality”, *Journal of Political Philosophy*, 18: 2, 137–155.
- J. Buchanan (1954) “Social Choice, Democracy, and Free Markets”, *Journal of Political Economy*, 62: 2, 114–123.
- J. Buchanan & G. Tullock (2004) *The Calculus of Consent. Logical Foundations of Constitutional Democracy*, Liberty Fund Inc.
- P. Casal (2007) “Why Sufficiency Is Not Enough”, *Ethics*, 117: 2, 296–326.

- A. Casella (2005) “Storable Votes”, *Games and Economic Behavior*, 51, 391–419.
- T. Christiano (1996) *The Rule of the Many: Fundamental Issues in Democratic Theory*, Oxford: Westview Press.
- T. Christiano (2008) “Democracy”, in Zalta, Edward N. (ed.), *The Stanford Encyclopedia of Philosophy*, <http://plato.stanford.edu/archives/fall2008/entries/democracy> (accessed 2013-08-18).
- J. Cohen (1986) “An Epistemic Conception of Democracy”, *Ethics*, 97: 1, 26-38.
- J. Cohen (2003) “Deliberation and Democratic Legitimacy”, in D. Matravers and J. Pike (eds.) *Debates in Contemporary Political Philosophy: An Anthology*, New York: Routledge.
- Condorcet, Marquis de (1785) *Essai sur l'application de l'analyse a la probabilité des décisions rendues 'a la pluralité de voix*, Paris: l'Imprimerie Royale, <http://gallica.bnf.fr/ark:/12148/bpt6k417181/f4.image> (accessed 2013-07-02).
- R. Crisp (2003) “Equality, Priority, and Compassion”, *Ethics*, 113: 4, 745–763.
- R. Crisp (2006) *Reasons and the Good*, Oxford: Clarendon Press.
- R. A. Dahl (1989) *Democracy and its Critics*, New Haven, CT: Yale University Press.
- F. Dietrich (2008) “The Premises of Condorcet's Jury Theorem Are Not Simultaneously Justified”, *Episteme*, 5: 1, 56–73.
- F. Dietrich and C. List (2004) “A Model of Jury Decisions Where All Jurors Have the Same Evidence”, *Synthese*, 142: 2, 175–202.
- F. Dietrich and K. Spiekermann (2010) “Independent Opinions?”, *Cahiers de Recherche de l'IHPST, Série "Décision, Rationalité, Interaction"*, Cahier DRI-2010-04.
- F. Dietrich and K. Spiekermann (2013) “Epistemic Democracy with Defensible Premises”, *Economics and Philosophy*, 29: 1, 87-120.
- G. Dimdins, H. Montgomery, and E. Norell (2011) “Egalitarian vs. Proportional Voting in Various Contexts: An Experimental Study”, *Oñati Socio-Legal Series*, 1: 5.
- A. Downs (1957) *An Economic Theory of Democracy*, New York: Harper & Row Publishers.
- J. S. Dryzek (2000) *Deliberative Democracy and Beyond*, Oxford: Oxford University Press.

- J. Duggan and T. Schwartz (2000) "Strategic Manipulability Without Resoluteness or Shared Beliefs: Gibbard-Satterthwaite Generalized", *Social Choice and Welfare*, 17, 85-93.
- M. Dummett (1984) *Voting Procedures*, Oxford: Clarendon Press.
- R. Dworkin (1987) "What is Equality? Part IV: Political Equality", *San Francisco Law Review*, 22: 1-30.
- J. Elster (2002) "The Market and the Forum: Three Varieties of Political Theory", in T. Christiano (ed.) *Philosophy and Democracy*, Oxford: Oxford University Press.
- D. M. Estlund (1989) "The Persistent Puzzle of the Minority Democrat", *American Philosophical Quarterly*, 26: 2, 143-150.
- D. M. Estlund (1990) "Democracy without Preference", *The Philosophical Review*, 99: 3, 397-423.
- D. M. Estlund (2008) *Democratic Authority: A Philosophical Framework*, Princeton: Princeton University Press.
- A. M. Feldman (1980) "Strongly Nonmanipulable Multi-Valued Collective Choice", *Public Choice*, 35: 4, 503-509.
- F. Feldman (1995) "Justice, Desert, and the Repugnant Conclusion", *Utilitas*, 7: 2, 189-206.
- F. Feldman (2004) *Pleasure and the Good Life*, Oxford: Clarendon Press.
- D. S. Felsenthal & M. Machover (2000) "Enlargement of the EU and Weighted Voting in Its Council Of Ministers", London: LSE Research Online, <http://eprints.lse.ac.uk/archive/00000407> (accessed 2013-08-18).
- M. Fleurbaey (*mimeo*) "One Stake One Vote", [http://www.ehu.es/p200-content/es/contenidos/informacion/00031\\_seminarios/es\\_00031\\_se/Fleurbaey\\_6-febrero.pdf](http://www.ehu.es/p200-content/es/contenidos/informacion/00031_seminarios/es_00031_se/Fleurbaey_6-febrero.pdf) (accessed 2013-08-18).
- H. Frankfurt (1987) "Equality as a Moral Ideal", *Ethics*, 98: 1, 21-43.
- M. Friedman (1994) "The Methodology of Positive Economics", in M. Martin and L. C. McIntyre (eds.) *Readings in the Philosophy of Social Science*, Cambridge, MA: Massachusetts Institute of Technology.
- A. Gibbard (1973) "Manipulation of Voting Schemes: A General Result", *Econometrica*, 41: 4, 587-601.
- C. Gould (1988) *Rethinking Democracy: Freedom and Social Cooperation in Politics, Economics and Society*, Cambridge: Cambridge University Press.
- K. Graham (1976) "Democracy, Paradox and the Real World", *Proceedings of the Aristotelian Society*, New Series, 76, 227-245.

- K. Graham (1996) "Voting and Motivation", *Analysis*, 56: 3, 184–190.
- J. Griffin (1988) *Well-being: Its Meaning, Measurement and Moral Importance*, Oxford: Clarendon Press.
- B. Grofman (1978) "Judgmental Competence of Individuals and Groups in a Dichotomous Choice Situation: Is a Majority of Heads Better Than One?", *The Journal of Mathematical Sociology*, 6: 1, 47-60.
- B. Grofman, G. Owen, S. Feld (1983) "Thirteen Theorems in Search of the Truth", *Theory and Decision* 15: 261–278.
- A. Gutmann and D. Thompson (1996) *Democracy and Disagreement*, Cambridge MA: Harvard University Press.
- R. Hardin (1993) "Public Choice Versus Democracy", in D. Copp, J. Hampton, J. E. Roemer (eds.) *The Idea of Democracy*, Cambridge: Cambridge University Press.
- R. Harrison (2005) "Democracy", in E. Craig (ed.) *Routledge Encyclopedia of Philosophy*, London: Routledge.
- J. Hawthorne (*mimeo*) "Voting in Search of the Public Good: The Probabilistic Logic of Majority Judgments", <http://faculty-staff.ou.edu/H/James.A.Hawthorne-1/Hawthorne--Jury-Theorems.pdf> (accessed 2013-08-18).
- D. Haybron (2008) *The Pursuit of Unhappiness*, Oxford: Clarendon Press.
- C. Heathwood (2010) "Welfare", in J. Skorupski (ed.) *Routledge Companion to Ethics*, London: Routledge.
- E. A. Henderson (2002) *Democracy and War: The End of an Illusion?*, Boulder CO: Lynne Rienner Publishers.
- D. Heyd and U. Segal (2006) "Democratically Elected Aristocracies", *Social Choice and Welfare*, 27, 103–127.
- R. Hortala-Vallve (2012) "Qualitative Voting", *Journal of Theoretical Politics*, 24, <http://jtp.sagepub.com/content/24/4/526> (accessed 2013-08-18).
- T. Hurka (1993) *Perfectionism*, Oxford: Oxford University Press.
- M. Jiborn and W. Rabinowicz (2003) "Reconsidering the Foole's Rejoinder: Backward Induction in Indefinitely Iterated Prisoner's Dilemmas", *Synthese*, 136: 2, 135-157.
- G. S. Kavka (1986) *Hobbesian Moral and Political Theory*, Princeton: Princeton University Press.

- L. A. Kornhauser and L. G. Sager (1986) "Unpacking the Court", *The Yale Law Journal*, 96: 1, 82-117.
- K. K. Ladha (1992) "The Condorcet Jury Theorem, Free Speech, and Correlated Votes", *American Journal of Political Science*, 36: 617-634.
- K. K. Ladha (1993) "Condorcet's Jury Theorem in Light of de Finetti's Theorem: Majority-Rule Voting with Correlated Votes", *Social Choice and Welfare*, 10, 69-85.
- D. Leech (2001) "Shareholder Voting Power and Corporate Governance: A Study of Large British Companies", *Nordic Journal of Political Economy*, 27, 33-54.
- C. List (2012) "Collective Wisdom: Lessons from the Theory of Judgment Aggregation", in H. Landemore and J. Elster (eds.) *Collective Wisdom: Principles and Mechanisms*, Cambridge: Cambridge University Press.
- C. List & R. E. Goodin (2001) "Epistemic Democracy: Generalizing the Condorcet Jury Theorem", *The Journal of Political Philosophy*, 9: 3, 277-306.
- G. Mackie (2003), *Democracy Defended*, Cambridge: Cambridge University Press.
- C. MacKinnon (1989) *Toward a Feminist Theory of the State*, Cambridge MA: Harvard University Press.
- J. Mansbridge (1990) "Expanding the Range of Formal Modeling", in J. Mansbridge (ed.) *Beyond Self-Interest*, Chicago: The University of Chicago Press.
- K. O. May (1952) "A Set of Independent and Sufficient Conditions for Simple Majority Decision", *Econometrica*, 20, 680-684.
- J. S. Mill (1861) *Considerations on Representative Government*, London: Saville and Edwards.
- J. S. Mill (1998) *Utilitarianism*, Oxford: Oxford University Press.
- N. R. Miller (1986) "Information, Electorates, and Democracy: Some Extensions and Interpretations of the Condorcet Jury Theorem", in: B. Grofman and G. Owen (eds.) *Information Pooling and Group Decision Making: Processings of the Second University of California, Irvine, Conference on Political Economy*, Greenwich CT.: JAI Press.
- A. Naess, J. A. Christophersen and K. Kvalø (1956) *Democracy, Ideology and Objectivity: Studies in the Semantics and Cognitive Analysis of Ideological Controversy*, Oslo : Oslo University Press.
- W. Nelson (1980) *On Justifying Democracy*, London: Routledge & Kegan Paul.
- S. Nitzan and J. Paroush (1982) "Optimal decision rules in uncertain dichotomous choice situations", *International Economic Review*, 23: 2, 289-297.
- R. Nozick (1974) *Anarchy, State, and Utopia*, Oxford : Blackwell.

- H. Nurmi (2010) "Voting Theory", in D. Rios Insua & S. French (eds.) *e-Democracy: A Group Decision and Negotiation Perspective*. Springer.
- M. Nussbaum (2000) *Women and Human Development: The Capabilities Approach*, Cambridge: Cambridge University Press.
- G. Owen, B. Grofman, S. Feld (1989) "Proving a Distribution-Free Generalization of the Condorcet Jury Theorem", *Mathematical Social Sciences* 17: 1–16.
- D. Parfit (1984) *Reasons and Persons*, Oxford: Clarendon Press.
- A. Petré (2007) "On Proportionality as a Democratic Ideal in Political Decision-Making" in F. Tersman (ed.) *The Viability and Desirability of Global Democracy*, Stockholm Studies in Democratic Theory 3, Stockholm: Stockholm University.
- P. Pettit (2001) "Deliberative Democracy and the Discursive Dilemma", *Philosophical Issues*, 11, 268-299.
- P. Pettit and R. Sugden (1989) "The Backward Induction Paradox", *The Journal of Philosophy*, 86: 169–82.
- A. Przeworski, M. E. Alvarez, J. A. Cheibub and F. Limongi (2000) *Democracy And Development: Political Institutions and Well-Being in the World, 1950–1990*, Cambridge: Cambridge University Press.
- W. Rabinowicz (*mimeo*) "Aggregation of Value Judgments vs. Aggregation of Preferences".
- D. W. Rae (1969) "Decision-Rules and Individual Values in Constitutional Choice", *The American Political Science Review*, 63: 1, 40-56.
- D. W. Rae and M. Taylor (1969) "Some Ambiguities in the Concept of 'Intensity'", *Polity*, 1: 3, 297–308.
- J. Rawls (1999) *A Theory of Justice*, revised edition, Oxford: Oxford University Press.
- W. Riker (1982) *Liberalism against Populism: A Confrontation Between the Theory of Democracy and the Theory of Social Choice*, San Francisco: W.H. Freeman & Company.
- J. Riley (1990) "Utilitarian Ethics and Democratic Government", *Ethics*, 100: 2, 335–348.
- A. Rosenberg (2008) *Philosophy of Social Science*, 3<sup>rd</sup> edition, Boulder CO: Westview Press.
- O. Rubin (2009) "The Merits of Democracy in Famine Protection – Fact or Fallacy?", *European Journal of Development Research*, 21, 699–717.

- M. A. Satterthwaite, (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.
- B. Saunders (2008) *Democracy-as-Fairness: Justice, Equal Chances and Lotteries*, D.Phil thesis in Politics in the Department of Politics and International Relations, University of Oxford, [http://www.academia.edu/148977/Democracy\\_as\\_Fairness](http://www.academia.edu/148977/Democracy_as_Fairness) (accessed 2013-08-18).
- J. Schumpeter (1975) *Capitalism, Socialism, and Democracy*, 3<sup>rd</sup> edition, New York: Harper & Row.
- A. Sen (1985) "Well-Being, Agency and Freedom: The Dewey Lectures 1984", *The Journal of Philosophy*, 82: 4, 169–221.
- A. Sen (1990) "Rational Fools: A Critique of the Behavioral Foundations of Economic Theory", in J. Mansbridge (ed.) *Beyond Self-Interest*, Chicago: The University of Chicago Press.
- A. Sen (1993) "Internal Consistency of Choice", *Econometrica*, 61: 3, 495–521.
- A. Sen (1995) "Rationality and Social Choice", *The American Economic Review*, 85: 1, 1–24.
- A. Sen (1999a) *Development as Freedom*, New York: Knopf.
- A. Sen (1999b) "The Possibility of Social Choice", *The American Economic Review*, 89: 3, 349-378.
- A. Sen (1997) "Description as Choice" in A. Sen, *Choice, Welfare and Measurement*, Cambridge MA: Harvard University Press
- A. Sen (1970) *Collective Choice and Social Welfare*, San Francisco: Holden-Day.
- L. Shapley and B. Grofman (1984) "Optimizing Group Judgmental Accuracy in the Presence of Interdependencies", *Public Choice*, 43: 329–43.
- L. Shields (2012) "The Prospects for Sufficiencyarianism", *Utilitas*, 24: 1, 101–117.
- P. Singer (1973) *Democracy and Disobedience*, Oxford: Clarendon Press.
- B. Skyrms (1998) "The Shadow of the Future", in J. L. Coleman and C. W. Morris (eds.) *Rational Commitment and Social Justice: Essays for Gregory Kavka*, Cambridge: Cambridge University Press.
- B. Skyrms (2001) "The Stag Hunt", *Proceedings and Addresses of the American Philosophical Association*, 75: 2, 31–41.
- J. H. Sobel (1993) "Backward-Induction Arguments: A Paradox Regained", *Philosophy of Science*, 60: 1, 114–133.

- S. S. Stevens (1946) "On the Theory of Scales of Measurement", *Science*, 103: 2684, 677–680.
- W. Sumner (1999) *Welfare, Happiness, and Ethics*, Oxford: Clarendon Press.
- T. Tännsjö (1992) *Populist Democracy: A Defence*, London: Routledge.
- T. Tännsjö (1998) *Hedonistic Utilitarianism*, Edinburgh: Edinburgh University Press.
- M. Taylor (1969) "Proof of a Theorem on Majority Rule", *Behavioral Science*, 14, 228–231.
- A. de Tocqueville (1990) *Democracy in America*, part I, New York: Vintage Books.
- J. Wolff (1994) "Democratic Voting and the Mixed-Motivation Problem", *Analysis*, 54: 4, 193-196.
- R. Wollheim (1962) "A paradox in the theory of democracy", in P. Laslett and W. G. Runciman (eds.) *Essays in Philosophy, Politics and Society*, New York: Barnes & Noble.
- H. P. Young (1988) "Condorcet's Theory of Voting", *American Political Science Review*, 82: 4, 1231–1244.