Katharina Berndt Rasmussen
katharina.berndt.rasmussen@iffs.se

The Rae-Taylor Theorem and the Weighted Majority Rule

Abstract: An influential theorem proposed by Rae (1969) and Taylor (1969) shows that the collective decision-rule which is individually optimal, for a constitution-maker behind a veil of ignorance, is the simple majority rule. A recent theorem by Brighouse and Fleurbaey (2010) and Fleurbaey (2008) shows that the weighted majority rule selects collectively optimal outcomes. In this paper, I argue that the Rae-Taylor framework contains hitherto underexplored resources that can be used to align their result with Brighouse and Fleurbaey's. More specifically, I argue that Rae's (1969) own discussion of the constitution-maker's possible biases points to a way of generalising his argument, which seamlessly transposes it to support the weighted, rather than the simple majority rule.

1 Introduction

An influential theorem by Douglas Rae[1] and Michael Taylor[2] shows that the individually optimal collective decision-rule, for a constitution-maker behind a veil of ignorance, is the *simple majority rule*. This rule, in binary collective decisions, assigns to every person an equal vote, and selects as outcome the option receiving more votes. A recent theorem by Harry Brighouse and Marc Fleurbaey[3] and by Fleurbaey[4] shows that the *weighted majority rule* selects collectively optimal outcomes. This rule, in binary collective decisions, assigns to every person a voting weight in proportion to her stakes and selects as outcome the option that receives more voting weights.

The weighted majority rule is extensionally equivalent to the simple majority rule, in decisions with equal stakes. However, since the former selects collectively optimal outcomes even when stakes are unequal, one would expect it to be optimal to choose for a constitution-maker under uncertainty. Thus, the Rae-Taylor theorem seems at odds with the Brighouse-Fleurbaey

[1] Rae (1969).
[2] Taylor (1969).
[3] Brighouse & Fleurbaey (2010).
[4] Fleurbaey (2008).

theorem. However, as Fleurbaey[5] points out, the latter is a generalisation of the former, which results from dropping Rae's assumption of equal intensities of preferences (i.e., of equal stakes).[6]

In this paper, I argue that the Rae-Taylor framework already contains the resources needed to derive the Brighouse-Fleurbaey result. More specifically, while Rae explicitly assumes equal intensities of preference, his own discussion of the constitution-maker's possible biases in effect allows for unequal intensities. We can describe this as an unacknowledged inconsistency in Rae's framework – or as an underexplored opening, pointing to a way of generalising his argument to support the weighted, rather than the simple majority rule.

Section 2 summarises Brighouse's and Fleurbaey's argument and spells out its implications for a veiled constitution-maker. Section 3 reconstructs Rae's argument, whose five controversial assumptions I point out in section 4. I show how they can be relaxed, along the lines already suggested by Rae, and that the resulting generalised argument advocates the weighted majority rule. Section 5 concludes.

2 The Brighouse-Fleurbaey argument

Assume that, for a given binary decision with options *x* and *y*, *x* yields a greater sum-total of well-being for the people it affects than *y*. Second, assume everyone either prefers *x* to *y* or *y* to *x* or is indifferent, and that the direction and intensity of her preference corresponds to her *stakes*, i.e., her difference in well-being units between the options. One way to understand the idea of stakes is that, for every decision, the voter's well-being level from the option that is worse for her constitutes a baseline, and her stakes equal the number of additional well-being units from the for her better option. Then those who prefer *x* to *y* do so with a greater overall intensity – and have a larger total amount of stakes – than those who prefer *y* to *x*. Third, assume a decision-rule assigning numbers of votes in proportion to stakes (assigning zero votes to indifferent individuals). Fourth, all *voters* (stakeholders with assigned votes) are

[5] Fleurbaey (2008).
[6] Cf. List (2013: §2.4); Conradt & List (2009: 730f.).

self-interested (seek to maximise their preference satisfaction), such that, in a vote between *x* and *y*, they choose their preferred option.[7] Then a decision-rule, which selects as outcome that option receiving the greater number of votes, selects the option with the greater sum-total of well-being. The rule described is the weighted majority rule.[8]

For an even number of votes, half of which support *x* and the other half *y*, the rule needs a tie-breaker. Ceteris paribus, any tie-breaker will do. Given that the votes are split evenly between the options, so are the stakes, which implies that the options are equal in sum-total of well-being. Thus, the weighted majority rule, complemented with any tie-breaker, selects the *collectively optimal outcome*, i.e., an option that is at least as high in sum-total of well-being as its alternative.

This argument provides a utilitarian social planner with a strong (prima facie) reason to implement the weighted majority rule.[9] What does it imply for the individual – as assumed, self-interested – voter? Does it provide her with a reason to accept this collective decision-rule? If this question concerns single instances of collective decision-making, the answer is no. The voter's acceptance depends on how the chips happen to fall: if she finds herself on the side of the minority-stakeholders, she has a (prima facie) self-interested reason to oppose the rule along with its outcome.

A more interesting question concerns the voter's second-order decision which rule to accept for any and all upcoming instances of collective decision-making. Which rule would be optimal for a voter *qua* constitution-maker, making this decision under ignorance? Intuitively, not knowing who she will be beyond a veil of ignorance, she should choose a rule maximising the sum-total of well-being for all. According to the above argument, this is the weighted majority rule.[10] A substantial argument for this conclusion runs as follows.

---

[7] The binary setup excludes strategic voting (Gibbard 1973); the question of the cost/rationality of voting (Downs 1957) is here disregarded entirely.

[8] Brighouse & Fleurbaey (2010), Fleurbaey (2008). They describe the rule as assigning voting *weights* to votes. Presenting it in terms of *numbers* of votes facilitates my arguments; though note that voters cannot split their votes between options.

[9] Its relevance is, admittedly, limited to a highly idealised context, with a planner assessing stakes and assigning votes correctly. For less idealised circumstances, other rules may approximate this ideal by letting each voter do the planner's task for herself; e.g., Hortala-Vallve's (2012) qualitative, or Casella's (2012) storable vote rule, or Tullock's (1970) logrolling practice.

[10] For an experimental study assessing the acceptance of weighted voting, see Montgomery and Dimdins (2009).

Katharina Berndt Rasmussen
katharina.berndt.rasmussen@iffs.se

Assume that the veil conceals which first-order decisions the constitution-maker will face and what her stakes will be. Second, the constitution-maker is *self-interested in a qualified sense*, seeking to maximise preference satisfaction while prioritising satisfaction of more intense preferences over satisfaction of less intense preferences. I.e., her normative criterion is to maximise intensity-weighted preference satisfaction. Then, not knowing whether she will be a majority- or minority-stakeholder, with high, low, or no stakes in any upcoming decision, she will satisfy this criterion by choosing a rule which, for every binary decision, assigns votes in proportion to stakes (corresponding to intensity-weighted preferences) and selects as outcome the option attracting an at least as great sum-total of votes as its alternative. I.e., she will choose the weighted majority rule. This conclusion provides the self-interested voter, *qua* constitution-maker, with a reason to accept this rule. Yet this runs counter to Rae's argument.

3 Rae's argument

Rae's constitution-maker is 'a single, anonymous individual who is self-interested in the sense of wishing 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'.[11]

Second, the constitution-maker chooses among n voting rules, for any group of $n≥3$ voters facing a binary decision between supporting policy x and defeating x (i.e., preserving status quo). At 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.[12] Between these extremes are the rules stating that x is passed only if $n–1$ ($n–2$; ...; or $n–(n–2)$) voters support it.

Third, the constitution-maker knows 'nothing about the (long-run) agenda which will confront the [collective], about the ways individuals will evaluate the proposals which do arise, or about

[11] Rae (1969), p. 41.
[12] Rae (1969), p. 49.

the factional structure of the [collective]'.[13] He just knows that each voter (including himself) will either support or reject each proposal, independently of any other. Rae's constitution-maker thus faces 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 (A) and (B), the 'values' expressed by the outcomes of the decision do not correspond to the constitution-maker's own. In (C) and (D), they do. His wish to 'have his way' as often as possible is precisified in Rae's individualist normative criterion: 'One should choose that decision-rule which minimizes the sum of the expected frequencies for (A) in which the [collective] does not impose a policy which his value schedule leads him to support, and (B) in which the [collective] imposes a policy which his value schedule leads him to oppose'.[14]

These assumptions generate a model within which the expected frequencies of (A) and (B) can be calculated for any voting rule that requires $n - m$ voters to vote for a policy in order to pass it. Rae assumes that $n > m \geq 0$. Under Rae's rule of consensus, the expected frequency of (A) is at its maximum: a policy is passed only if everyone supports it. The expected frequency of (B), however, is zero: the constitution-maker's rejection will suffice to defeat a policy. Under Rae's rule of individual initiative, the tables are turned. The expected frequency of (A) is zero: the constitution-maker's support will suffice for a policy to pass. The expected frequency of (B) is, however, quite high: a policy is defeated only if no one supports it. Between these two extremes, the expected frequencies for (A) are monotonically increasing with the number of individuals whose support is required for a policy to pass, while the expected frequencies for (B) are monotonically decreasing. The frequency curves are thus opposed.

[13] Rae (1969), p. 41.
[14] Rae (1969), p. 42.

Rae's normative criterion requires that the *sum* of the expected frequencies for (A) and (B) is minimised. Rae illustrates that this minimum is located between the two extremes, rule of consensus and rule of individual initiative. For an odd number of voters, 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+1)/2.

Hence, simple majority rule, requiring that more than half the voters support a policy for it to pass, is an optimal decision-rule. Rae concludes that 'majority-rule is as good (i.e. optimal) as any alternative decision-rule, given the model proposed here'.[15] In the long run, it ensures the constitution-maker's preferred outcomes as often as possible. This provides the constitution-maker with a reason to accept this rule.

4 Adjusting Rae's argument

I now examine five central restrictions of Rae's argument: (i) Rae's n decision-rules "exhaust the available alternatives", such that, e.g., weighted decision-rules are excluded from the outset.[16] (ii) For every policy x and every voter i, i either supports or rejects x, leaving no room for indifference, i.e., ranking x as just as good as the status quo. (iii) Framing the argument in terms of supporting versus rejecting policies makes all decisions status-quo dependent, leaving aside decisions between two options, neither of which is the status quo, (e.g., two mutually exclusive policies unanimously ranked above status quo). (iv) Rae explicitly disregards as theoretically intractable the 'problem of intensity', i.e., the idea that in some decisions there might be more at stake for some voters than for others.[17] (v) The constitution-maker is not biased in favour of one of the options, ruling out, e.g., a conservative bias in

[15] Rae (1969), p. 52. Strictly speaking, Rae shows that a "non-minority rule", requiring that *at least half* the voters support a policy in order for it to pass, is optimal. However, for the special case with a policy supported by exactly n/2 voters, the constitution-maker is indifferent between the simple majority rule (policy not passed) and the non-minority rule (policy passed), since he has an equal chance of being among either half of the voters. Thus, either rule is equally as good for him.
[16] Rae (1969), p. 40.
[17] Rae (1969), p. 41, footnote 6. Cf. Rae and Taylor (1969).

favour of preserving the status quo. Rae characterises such bias as a '*positional* (as opposed to substantive) *preference*'.[18]

Undoubtedly, Rae's argument is sound given these restrictions; and uncontroversially, we may make it more general by suitably relaxing them, thereby possibly modifying its conclusion. The next section, however, highlights how Rae's own considerations point us into the direction of such a generalised argument. Starting from (v), I show that relaxing the no-bias assumption, in accordance with Rae's suggestions, paves the way for relaxing the others.

4.1 The problem of bias

Rae considers a constitution-maker with a general conservative bias in favour of the status quo, assigning more disvalue to (B) than to (A). His normative criterion now needs adjustment: the optimal decision rule minimises the sum of the *weighted* expected frequencies of (A) and (B), with the weights chosen in proportion to the assigned (dis)values.[19]

Rae discusses the weights of 1 and 2, respectively, for (A) and (B). Intuitively, the idea is that bad action (B) is twice as bad as bad inaction (A). Stressing 'the enormous difficulty of supplying meaningful quantities for [these weights]', Rae concedes that, '[o]n the assumption that the weights themselves make sense', the adjusted normative criterion singles out another optimal voting rule: the two-thirds majority rule, according to which a policy is passed only if 2/3 of the *n* voters support it.[20]

We should pause to note the extent to which this concession changes Rae's model. Originally, the constitution-maker ranks the events according to only one factor: correspondence between preference and outcome – (C) or (D) – is ranked above non-correspondence – (A) or (B). By introducing a conservative bias, Rae introduces an additional factor, ranking non-preferred status quo – (A) – above non-preferred change – (B). Thus, the constitution-makers now ranks the two options he values – (C) or (D) – over one he disvalues – (A) – over one he

[18] Rae (1969), p. 52.
[19] Rae (1969), p. 52-53.
[20] Rae (1969), p. 53.

disvalues *even more* – (B). Hence, given that he supports a policy, he prefers (D) to (A), and given that he opposes a policy, he prefers (C) to (B), and the latter preference's intensity is *greater* than that of the former. Relaxing assumption (v) by introducing bias thus implies introducing 'intensity', i.e., varying stakes. This amounts to relaxing assumption (iv). Rae's '*positional*' preference is thus not only opposed to substantive preference, but also to the binary (on-off) picture of preference satisfaction underlying the whole of Rae's model.

In dealing with bias, Rae thus points out a way to deal with the problem of intensities. His specific suggestion holds for a constitution-maker who is always and only biased in one way — in this case, conservatively. For generality, we should allow that he may be conservatively biased only in certain decision, e.g., concerning family issues. In decisions concerning, e.g., education and the sciences, he may instead have an anti-conservative (innovative) bias, or no bias at all. Yet for such a voter, Rae's adjusted normative criterion with fixed weights becomes irrelevant.

An upshot of generalising the assumption of varying biases to one concerning varying stakes is that we do not have to frame the events facing the constitution-maker in terms of *passed* or *defeated* policy *x* (both having the status quo non-*x* as their baseline). We can frame them instead in terms of whether or not the collective ranking of any two options *x* and *y* corresponds to the constitution-maker's individual ranking of these options. This amounts to relaxing assumption (iii) and reduces the number of events the constitution-maker has to take into consideration:

> (I) *Correspondence*: The constitution-maker ranks *x* above *y*, as does the collective.
>
> (II) *Non-correspondence*: The constitution-maker ranks *x* above *y*, but the collective does not.

(I) is equivalent to the union of (C) and (D). (II) is equivalent to the union of (A), (B), and additional event (E). (E) covers all cases of non-correspondence, such that the constitution-maker is not indifferent between x and y, while the collective ranking is indifferent. This possibility was missing in Rae's model.

Within this framework, we can also account for another missing event:

> (III) *Individual indifference*: The constitution-maker is indifferent between *x* and *y*, while the collective either ranks one above the other or is indifferent as well.

Thus, relaxing (v), the no-bias assumption, helps reframe the constitution-maker's decision problem in a way which ultimately allows us to relax even assumption (ii).

In the next section, I suggest a normative criterion for a self-interested constitution-maker considering the events of (I), (II), and (III) in this less restricted context. I then show how the criterion, along Rae's own lines of reasoning, seamlessly leads him to accept the weighted majority rule. This shows how even assumption (i), the restriction to Rae's set of *n* voting rules, can be relaxed.

4.2 Deriving the weighted majority rule from the Raean framework

Considering (I), (II), and (III), the constitution-maker wants as much correspondence and as little non-correspondence as possible. Indifference event (III) does not matter to him (unless his goal of 'having his way' is a fetish). However, not all correspondence is equally good and not all non-correspondence equally bad for him, due to varying stakes. Accordingly, he wants as much correspondence as possible, especially in high-stakes decisions, and as little non-correspondence as possible, again, especially in high-stakes decisions.

To precisify this idea, the constitution-maker wants to maximise the sum of the weighted expected frequency of (I) minus the weighted expected frequency of (II), with weights chosen in proportion to his stakes. This *stakes-sensitive normative criterion* applies to equal and unequal stakes-cases alike. (We can see from its formulation why (III) becomes irrelevant: this event occurs only when the constitution-maker is indifferent, such that his stakes, and hence weights, will be zero. Thus, including (III) makes no difference.)

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 (I) and (II). But it helps if the constitution-maker shifts perspective. He knows that, in each upcoming decision, he will have some number of stakes, 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. What he wants can now be re-described as having as many of his stakes in instances of (I) and as few in instances of (II) as possible. Thus, for every one of his stakes, he wants to maximise the probability that it 'occurs' in (I), or — equivalently — minimise the probability that it 'occurs' in (II). (That a voter's stake 'occurs' in (I) (or (II)) simply means that it is one of his stakes in a decision where the for him better (worse) option is selected.) Thus, he wishes to minimise the expected frequency of (II) for each of his stakes, rather than for himself. This is the *stakes-centred normative criterion*.

I now assimilate this criterion to Rae's original criterion, to facilitate the assimilation of his argument to my purposes. As stated, (II) is equivalent to the union of (A), (B), and (E). Rae's criterion only considers (A) and (B). Assume, initially, that there are only decisions with an odd total number of stakes, such that there cannot be any collective indifference. Then, (E) can be disregarded, such that (II) is equivalent to the union of (A) and (B). 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 stakes-centred normative criterion, except that it focuses on the constitution-maker — as one of *n* voters — instead of on any one of his stakes — as one among a total of *s* stakes.

Rae shows that the optimal decision rule, according to his criterion, requires that to get *x* passed, there are $(n+1)/2$ voters supporting *x*, for an odd number of voters. Along the same lines, we can now conclude that the optimal decision rule, according to the stakes-centred normative criterion, requires that to get *x* passed, there are $(s+1)/2$ stakes 'supporting' *x*, for an odd number of stakes. Recall that we defined a voter's stakes as equalling her number of additional well-being units from the for her better option, compared to the baseline level of the for her worse option. Each stake can thus be said to 'support' the for her better option. Thus, the support of a simple majority of stakes is required for a policy to be selected. This is ensured by the weighted majority rule, which assigns numbers of votes in proportion to stakes and selects as outcome the option that gets a simple majority of votes.

Assume now that there are decisions with an even total number of stakes. Then there are three possibilities. One, there is no collective indifference, i.e., there is a majority of stakes in favour of one of the options. Then, the above argument holds even here. Two, there is collective indifference between the options — and the constitution-maker is indifferent as well. Then, he does not care which voting rule is employed. The weighted majority rule, with any tie-breaker, is as good for him as any other. Three, there is collective indifference between options *x* and *y* — and the constitution-maker is *not* indifferent. This is event (E). Since there are as many stakes in total for *x* as there are for *y*, the constitution-maker knows that it is equally likely that his stakes will support either option. Hence, the weighted majority rule, with any tie-breaker choosing either one of the options, is as good for him as any other.

The weighted majority rule is thus better for the constitution-maker than any other, in odd-stakes cases as well as even-stakes cases without collective indifference. And it is as good for him as any other, in even-stakes cases with collective indifference. He will thus choose it behind his veil of ignorance.[21]

5 Conclusion

I have shown that the Rae-Taylor theorem can be taken a good step further, and that Rae's own discussion already contains the resources necessary to generalise his result. Rae's consideration of the problem of bias points out a way to handle varying stakes, to allow status-quo independence for describing the options, and to allow voter indifference. The resulting generalised argument supports the weighted majority rule, as the individually optimal collective decision-rule, and thus aligns with the Brighouse-Fleurbaey theorem's implication for individual optimality.[22]

[21] Replace 'stakes' with 'constituents', and you get a constitution-maker version of Barberá and Jackson's (2006) argument for a weighted majority rule for political representatives, with weights proportional to their number of constituents.

[22] Acknowledgement: [removed for peer review].

References

Barberá, S. and M. O. Jackson. 2006. “On the Weights of Nations: Assigning Voting Weights in a Heterogeneous Union”. *Journal of Political Economy* 114: 2, 317-339.

Brighouse, H. and M. Fleurbaey. 2010. “Democracy and Proportionality”. *Journal of Political Philosophy* 18: 2, 137–155.

Casella, A. 2012. *Storable Votes: Protecting the Minority Voice*. Oxford: Oxford University Press.

Conradt, L. and C. List. 2009. “Group Decisions in Humans and Animals: A Survey”. *Philosophical Transactions of the Royal Society: Biological Sciences* 364: 1518, 719-742.

Downs, A. 1957. *An Economic Theory of Democracy*. New York: Harper & Row Publishers.

Fleurbaey, M. 2008. “One Stake One Vote”. https://www.ehu.eus/documents/1619441/1670503/stake.pdf.

Gibbard, A. 1973. “Manipulation of Voting Schemes: A General Result”. *Econometrica* 41: 4, 587-601.

Hortala-Vallve, R. 2012. “Qualitative voting”. *Journal of Theoretical Politics* 24: 4, 526–554.

Montgomery, H. and G. Dimdins. 2009. “Negative and positive stakes in plural voting: An experimental study”. *Baltic Journal of Psychology*. 10: 1-2, 14-24.

List, C. 2013. "Social Choice Theory". *The Stanford Encyclopedia of Philosophy* (Winter 2013 Edition), Edward N. Zalta (ed.), <https://plato.stanford.edu/archives/win2013/entries/social-choice/>.

Rae, D. W. 1969. “Decision-Rules and Individual Values in Constitutional Choice”. *The American Political Science Review* 63: 1, 40-56.

Rae, D. W. and M. Taylor. 1969. “Some Ambiguities in the Concept of 'Intensity”. *Polity* 1: 3, 297–308.

Taylor, M. 1969. “Proof of a Theorem on Majority Rule”. *Behavioral Science* 14, 228–231.

Tullock, G. 1970 “A Simple Algebraic Logrolling Model”. *The American Economic Review* 60: 3, 419-426.