



# Fair group decisions via non-deterministic proportional consensus

Jobst Heitzig<sup>1</sup> · Forest W. Simmons<sup>2</sup> · Sara M. Constantino<sup>3,4,5</sup>

Received: 15 May 2022 / Accepted: 5 April 2024  
© The Author(s) 2024

## Abstract

Are there group decision methods which (i) give everyone, including minorities, an equal share of effective decision power even when voters act strategically, (ii) promote consensus and equality, rather than polarization and inequality, and (iii) do not favour the status quo or rely too much on chance? We describe two non-deterministic group decision methods that meet these criteria, one based on automatic bargaining over lotteries, the other on conditional commitments to approve compromise options. Using theoretical analysis, agent-based simulations and a behavioral experiment, we show that these methods prevent majorities from consistently suppressing minorities, which can happen in deterministic methods, and keeps proponents of the status quo from blocking decisions, as in other consensus-based approaches. Our simulations show that these methods achieve aggregate welfare comparable to common voting methods, while employing chance judiciously, and that the welfare costs of fairness and consensus are small compared to the inequality costs of majoritarianism. In an incentivized experiment with naive participants, we find that a sizable fraction of participants prefers to use a non-deterministic voting method over Plurality Voting to allocate monetary resources. However, this depends critically on their position within the group. Those in the majority show a strong preference for majoritarian voting methods.

---

✉ Jobst Heitzig  
heitzig@pik-potsdam.de

- <sup>1</sup> FutureLab on Game Theory and Networks of Interacting Agents, Potsdam Institute for Climate Impact Research, PO Box 60 12 03, 14412 Potsdam, Germany
- <sup>2</sup> Liberal Arts and Mathematics Division, Portland Community College, Cascade Campus, 705 N. Killingsworth Street, Portland, OR 97217, USA
- <sup>3</sup> School of Public Policy and Urban Affairs, Northeastern University, 310 Renaissance Park, Boston, MA 02115, USA
- <sup>4</sup> Department of Psychology, Northeastern University, 105 Forsyth St, Boston, MA 02115, USA
- <sup>5</sup> School of Public and International Affairs, Princeton University, 20 Prospect Ave, Princeton, NJ 08540, USA

## 1 Introduction

Groups must often reach important collective decisions, despite diverse and sometimes contested views on the best course of action. For example, city council members decide on which public services to invest in and how to allocate them, board members can weigh in on hiring and other company initiatives, and community members may vote on how to allocate budgets in places with participatory budgeting. Whether informally or formally instituted through a voting mechanism, group decisions often employ some form of majority rule to reach a decision. This is true of Plurality Voting, but also other common voting methods, such as Approval Voting (Brams and Fishburn 1978), which are designed to identify “the will of the majority”. But majority rule, perceived as a cornerstone of democracy, can also oppress minorities (Lewis 2013). At a macro scale, this ‘tyranny of the majority’ may lead to separatism or violent conflict (Collier 2004; Cederman et al. 2010) and ultimately to welfare losses. To give just one example, Devotta (2005) documents the role of majority voting in the Sri Lankan separatist war, where the political structure led to ethnic outbidding, which resulted in ethnic conflict with the Tamil minority.

One approach to ensuring that minority positions receive adequate representation is to develop a voting mechanism that takes into account the fundamental fairness principle of proportionality (Cohen 1997; Cederman et al. 2010). Proportional representation does this to some extent, but if it is only used for the election of a representative body, which then itself uses a ‘standard’ (majoritarian) voting mechanism, the tyranny of the majority may be upheld (Zakaria 1997). This is because proportional representation does not imply proportional power: even a 49 percent faction may not have influence over decisions. For example, the strong and increasing polarization in the US Senate over the past several decades (McCarty et al. 2016) has limited the effectiveness of the minority party even though it has typically held more than 40% of the seats. This effect can be quantitatively captured by the Banzhaf and Shapley–Shubik power indices (Dubey and Shapley 1979).

So how can effective decision-making power be proportionally distributed? In this paper, we use a mixed-methods approach to provide an in-depth analysis of two group decision methods—including a novel method that uses conditional commitments to support potential consensus options that depend on others’ support. These methods achieve fairness by distributing power proportionally and increase welfare efficiency by supporting not just full but also partial consensus and compromise. We discuss these methods in detail below.

Smaller groups often try to overcome the majority problem through deliberation that is aimed at consensus or consent, but this can be time-consuming and difficult in strategic contexts or in larger group decision settings (Davis 1992). Deliberation may also be perceived as less legitimate than formal voting (Persson et al. 2013). Furthermore, supporters of the status quo may block consensus indefinitely. If the group decision protocol includes a fallback method that is applied if no consensus is reached by some deadline, and if this fallback method is majoritarian, then the majority has an incentive to simply wait until the fallback method is invoked. Hence common consensus procedures often favor the status quo (Bouton et al. 2018) or are effectively majoritarian. If voters can act strategically, then most common group decision methods

are also effectively majoritarian. Note that in public decision-making, strategic voting behavior appears to be common (Bouton 2013; Kawai and Watanabe 2013; Spenkuch et al. 2018).

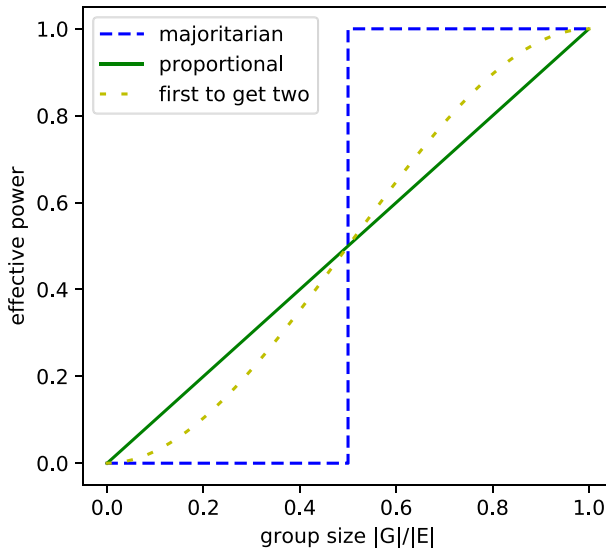
Within accepted or dominant social choice mechanisms, there is thus limited ability to give voice to minority positions (May 1952). Addressing this dimension may require re-considering certain features of standard social choice mechanisms, such as the role of chance or the distribution of voters' weights in preference aggregation. Deterministic decision methods that use chance only to resolve ties tend to suppress minority positions unless voters' weights are adjusted over time in response to their satisfaction to achieve some form of long-term proportionality, as in "Perpetual Voting" (Lackner 2020). However, it is relatively simple to distribute effective power proportionally with *non-deterministic* methods. In such methods, the winning option is sometimes determined by some amount of chance, not only to resolve ties, but also to achieve fairness or provide incentives for cooperation. For non-deterministic methods, it is natural to measure the effective power of a group by the amount of winning probability that the group can guarantee their chosen option. Methods under which any majority of  $> 50\%$  can guarantee their chosen option 100% winning probability, such as Plurality Voting, Approval Voting, Range Voting, Instant Runoff Voting, and Simpson–Kramer Condorcet, will be called 'majoritarian' here.

For example, the 'Random Ballot' (aka 'Lottery Voting') method (Amar 1984), despite having many undesirable properties, does manage to distribute effective power in a proportional way. In this method, voters indicate a single option on their ballot, and then a single ballot is drawn at random to decide the winner. Variants of this method can also be used to achieve a distribution of effective power that falls somewhere between the "majority-takes-all" approach of majoritarian methods and a perfectly proportional distribution. For example, one could draw a sequence of standard ballots until one option's vote count is two. This method, which we can refer to as "first to get two", would lead to the S-shaped power distribution shown in Fig. 1.<sup>1</sup>

The question of how to support consensus while also distributing power proportionally is more challenging. However, it is also important to ensure that the group decision method leads to fair and also efficient outcomes that avoid extreme results and foster social cohesion. 'Random Ballot' does the opposite: voters have no incentive to select a potential consensus option, even if it was everybody's second-best choice and a good compromise. Recent theoretical results show that the combination of fairness (by which we mean proportional allocation of power) and efficiency (by which we mean selecting good compromise options that lead to higher welfare) can be achieved if chance is used to incentivize consensus (Heitzig and Simmons 2012; Börgers and Smith 2014). In those papers, the main innovation was to use a lottery of voters' first choices as the fall-back to a consensus vote.

While employing chance to incentivize consensus can achieve fairness and efficiency criteria, introducing uncertainty in collective decision-making may also be aversive to people (Gill and Gaiouris 2002). However, real-world problems typically involve unavoidable stochastic risk and other forms of uncertainty, which is thus a fea-

<sup>1</sup> While it might seem that the area under that power curve must be  $1/2$  for all possible methods, this is not so: With the Borda method, the curve is a step function jumping from 0 to 1 at  $|G|/|E| = 2/3$ . *Supplement 3.1.4* contains derivations of these facts.

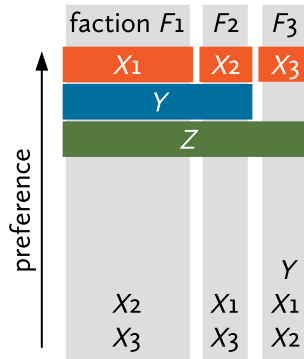


**Fig. 1** Distribution of effective power by type of group decision method. Effective power of a subgroup  $G$  of an electorate  $E$ , as a function of relative group size  $|G|/|E|$ , for different types of voting methods. Dashed: Majoritarian, deterministic methods. Solid: Proportional, non-deterministic methods such as Random Ballot, Full Consensus/Random Ballot, Full Consensus/Random Ballot/Ratings, and two novel methods: Nash Lottery and MaxParC. Dotted: An example of a non-proportional, non-deterministic method (“first to get two”) (color figure online)

ture of everyday life that constituents are familiar with (Carnap 1947). Additionally, there is prominent theoretical literature on such non-deterministic collective decision methods (e.g., Brandl et al. 2016). Furthermore, non-deterministic procedures are routinely used in diverse contexts, including learning (Cross 1973), optimization (Kingma and Ba 2014), strategic interactions (Harsanyi 1973), tax audits and mechanism design more generally, or the allocation of indivisible resources (e.g. school choice) (Trojan 2012). There are also an increasing number of proposals to use these methods for deciding on the composition of citizens’ councils (Flanigan et al. 2021) or even appointing officers, as in ancient Athens. These examples demonstrate that introducing an element of chance is one way to achieve certain desirable properties of collective decision mechanisms.

**Problem statement.** In this article, we adopt the working hypothesis that non-deterministic voting methods with clear advantages in terms of ethical criteria (e.g. fairness) or economic criteria (e.g. welfare) should be considered by constituents in situations where (a) collective decisions are made often and (b) the decisions are not momentous. This includes regularly occurring group decisions about the governance and management of day-to-day affairs, and excludes large-scale and infrequent elections of representatives.

With this context in mind, we provide a normative analysis of two such group decision methods, introduced in Sect. 2, and show that they achieve fairness by distributing power proportionally and increase welfare efficiency by supporting not just



**Fig. 2** Archetypal group decision problem with potential for suppression of minorities, partial consensus, or full consensus. Consider a situation where each of three factions of different size (column width) has a unique favourite outcome (topmost). However, there might also be a potential ‘partial consensus’ option  $Y$  and/or a potential ‘full consensus’  $Z$ . Assuming strategic voters, majoritarian methods will result in  $X_1$  for sure if faction  $F_1$  forms a majority. A notable exception is the Borda method, which might result in  $Y$  or  $Z$  unless  $F_1$  forms at least a two-thirds majority. In contrast, the two non-deterministic methods we consider—*Nash Lottery* and *MaxParC*—will select  $Z$  with certainty when it is present (green); one of  $X_{1,2,3}$  with probabilities proportional to faction size when neither  $Y$  nor  $Z$  are present (orange); and  $Y$  or  $X_3$  with proportional probabilities if  $Y$  but not  $Z$  is present (blue) (color figure online)

full but also partial consensus and compromise (Sect. 3). One of these methods, the *Nash Lottery*, is adapted from a different context and can be interpreted as a form of automatic bargaining using the Nash bargaining solution. The other, *Maximal Partial Consensus (MaxParC)*, is novel and is based on the idea of making one’s commitment to support a potential consensus option conditional on others’ support for that option. We complement this theoretical analysis with two empirical approaches: one based on large-scale experiments with simulated agents (Sect. 4) and one based on a virtual lab experiment with real participants recruited from Amazon Mechanical Turk (Sect. 5). The simulation analyses allow us to assess any potential welfare losses from implementing these decision rules relative to common alternatives and how agent and problem characteristics affect the amount of randomness needed to reach a group decision. In the virtual lab experiment, we elicit participants’ preferences between MaxParC vs. Plurality Voting, and assess whether preferences depend on the degree of distributive inequality facing the group or demographic and socioeconomic characteristics of our respondents. Given the complexity of this mixed-methods approach, we have complemented the main text with an extensive *Supplement* containing more details about the three parts of the paper (theory, simulations, behavioural experiment), but have still tried to keep the main text largely self-contained.

**Example.** As a paradigmatic test case (Fig. 2), consider a hypothetical group of three factions,  $F_1, F_2, F_3$ , with sizes  $S_{1,2,3}$  (in percent of voters). Assume that each faction has a preferred option,  $X_{1,2,3}$ , respectively, that is not liked by the other two factions. Now suppose there is a fourth option,  $Y$ , which is not liked by faction  $F_3$ , but which is liked by factions  $F_{1,2}$  almost as much as their respective favourites,  $X_{1,2}$ . Let us call  $Y$  the ‘partial consensus’ outcome for factions  $F_{1,2}$ . While what one might call

‘welfare efficiency’ (picking an option that is scoring high in terms of some metric of utility or welfare) seems to require that option  $Y$  has a considerable chance of being selected, proportionality requires that  $X_3$  also has some chance of winning—namely, an  $S_3\%$  chance. The two group decision methods described in this paper assign winning probabilities of  $S_1 + S_2\%$  to  $Y$  and  $S_3\%$  to  $X_3$ . Importantly, they do so not only if all voters vote “sincerely” (i.e., honestly according to their true preferences) but also if some or all voters vote strategically (e.g., by misrepresenting their preferences in a certain way).

If we now introduce a fifth option,  $Z$ , that  $F_{1,2}$  like only slightly less than  $Y$ , and which faction  $F_3$  likes almost as much as  $X_3$ , then both decision methods will select this ‘full consensus’ option  $Z$  for sure. In contrast, if the faction sizes fulfil  $S_1 > S_2 + S_3$  and if voters act strategically, then virtually all existing group decision methods will either pick  $X_1$  with certainty, or will assign probabilities of  $S_{1,2,3}\%$  to  $X_{1,2,3}$ , respectively. In both cases, these methods would ignore the potential compromise options  $Y$  and  $Z$  and would thus achieve lower overall welfare as compared to the two non-deterministic methods highlighted in this paper. In fact, no deterministic decision method is able to ensure that  $Y$  gets a chance without also rendering faction  $F_3$ ’s votes completely irrelevant! Recognizing  $F_3$ ’s votes can only be achieved through the introduction of a judicious amount of chance.

So how exactly should a non-deterministic group decision method be designed to achieve both efficiency and fairness (taken here to mean effective proportionality), even when voters act strategically, but also other consistency requirements typically studied in social choice theory—such as anonymity, neutrality, monotonicity, and clone-proofness—that make it plausible and hard to manipulate? We demonstrate two such methods using a combination of ingredients from existing group decision methods (i.e., Approval Voting and Random Ballot), game theoretical concepts (i.e., the Nash Bargaining Solution), and Granovetter’s “threshold” model of social mobilisation (Granovetter 1978).

## 2 Two possible solutions

**The Nash Lottery.** The first method we highlight and include in our study is what we call the *Nash Lottery* (*NL*). It is an adaptation of ‘Nash Max Product’ or ‘Maximum Nash Welfare’ from the literature on fair division of resources. As suggested in Aziz et al. (2019), we translate it to our context by interpreting *winning probability* as a “resource” to be divided fairly, and study the strategic implications of doing so. The Nash Lottery can be interpreted as a form of *automatic bargaining* by means of the well-known Nash Bargaining Solution. Similar to score-based methods such as Range Voting (RV) (Laslier and Sanver 2010), it asks each voter,  $i$ , to give a *rating*,  $0 \leq r_{ix} \leq 100$ , for each option  $x$ . Like Range Voting, the Nash Lottery then assigns winning probabilities,  $p_x$ , to all options  $x$  so that a certain objective quantity,  $f(r, p)$ , is maximized.

Range Voting maximizes the quantity  $f(r, p) = \sum_i \sum_x r_{ix} p_x$ , which is motivated by its formal similarity to a utilitarian welfare function. This results in a majoritarian method that is deterministic (i.e., we usually have  $p_x = 1$  for some  $x$  except for ties).

This determinism is because  $f(r, p)$  is a linear function of  $p$ . But that method neither distributes power proportionally nor supports consensus when voters are strategic. In the example of Fig. 2, faction  $F_1$  will quickly notice they can make  $X_1$  win for sure by putting  $r_{ix} = 0$  for all other options. Indeed, Range Voting is more or less strategically equivalent to the simpler Approval Voting (Dellis 2010). As a consequence, strategic voters almost never have an incentive to make use of any other rating than 0 or 100. In case there is no Condorcet winner (an option preferred to each other option by some majority), which is not too unlikely (Jones et al. 1995), there is not even any strategic equilibrium between factions, and thus the outcome is largely unpredictable.

The Nash Lottery instead maximizes the quantity

$$f(r, p) = \sum_i \ln \left( \sum_x r_{ix} p_x \right), \quad (1)$$

which gives a non-deterministic method (i.e., usually several options have a positive winning probability  $p_x$ ) that supports both full and partial consensus. In *Supplement 3.1.5*, we prove that in situations similar to Fig. 2, a full consensus will be the sure winner under the Nash Lottery, and a partial consensus would get a proportional share of the winning probability. This would be so both in the case where voters are sincere and in a certain strategic equilibrium. Even more so, we show in *Supplement 3.1.4* that using the logarithm rather than any other function of  $\sum_x r_{ix} p_x$  is the only possible way to achieve a proportional distribution of effective power.

The Nash Lottery is conceptually simple, and we will see below that it has some additional desirable properties beyond distributing power proportionally and incentivizing consensus, such as being immune to certain manipulations, e.g., adding a bad option to affect the tallying process in ways beneficial to the manipulator. However, this group decision method also has three properties we consider important drawbacks. Its tallying procedure requires performing an optimization task for which we know of no pen-and-paper solution method. It also lacks an intuitive ‘monotonicity’ property: as we will prove below, when a new option is added or a voter increases some existing option’s rating, this might increase rather than decrease some other option’s winning probability. Last but not least, we will see in our simulations that the Nash Lottery employs considerably more randomness than our second method: it less often produces a deterministic result, and the resulting lotteries have larger entropy and smaller maximal probability.

**Maximal Partial Consensus (MaxParC).** All three issues are addressed by our second method that we develop in this paper, *Maximal Partial Consensus (MaxParC)*. This method is conceptually more complex, but is also strongly monotonic, easier to tally, and leads more often to a deterministic winner in simulations. Like with Random Ballot, each voter’s “vote” represents an equal share,  $1/N$ , of the winning probability, where  $N$  is the number of voters. But unlike Random Ballot, MaxParC allows each voter to “safely” transfer their vote from their favourite option to a potential consensus option, thereby incentivising voters to find and implement good compromise options. This safety is achieved by making sure the vote transfer only becomes effective if enough other voters transfer their votes as well. This effect is achieved by using



a design element that allows voters to make *conditional commitments to approve compromise options*.

To motivate the need for conditional commitments and our precise design, we can first consider a simpler method that distributes power proportionally but does not sufficiently incentivise compromise: the ‘Conditional Utilitarian Rule’ (CUR) from Duddy (2015) and Aziz et al. (2019), a probabilistic, proportional method based on approval ballots. In CUR, each voter states whether or not they *approve* each option. A ballot is then drawn at random and, from the options approved on this ballot, the one with the largest overall number of approving ballots wins.

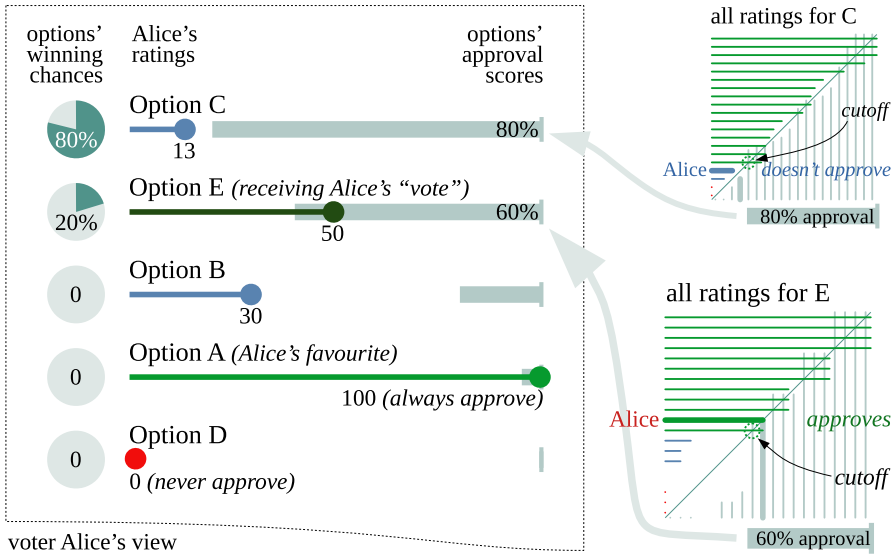
In the test case of Fig. 2 with  $S_1 > S_2 + S_3$ , if all voters approved  $Z$ ,  $Z$  would be the most-approved option and CUR would elect  $Z$  for sure. But faction  $F_2$  would soon realize that if they all withdrew their approval of  $Z$ ,  $Z$  would still be the most-approved option. All votes from  $F_1 \cup F_3$  would still go to  $Z$ , while the votes from  $F_2$  would now go to their favourite,  $X_2$ , which is a strict improvement for  $F_2$ . Approving  $Z$ , and thus transferring votes from  $X_1$  to  $Z$ , can be seen as an “unsafe” contribution by  $F_1$  to a public good that can be exploited by  $F_2$  who can “free ride” by not approving  $Z$ . As in a public goods game, unanimous approval of the potential consensus option is not a strategic equilibrium under CUR, hence  $Z$  would not necessarily win.

The solution we propose is that  $F_1 \cup F_3$  should be given the possibility to condition their approval of  $Z$  on others’ approval of  $Z$ . This would mean that when  $F_2$  withdraws approval of  $Z$ , others do so as well, resulting in everyone’s votes going back to their favorite options  $X_i$ , which is no longer an improvement for the deviators in  $F_2$ . While one can imagine various ways by which an approval might be made conditional, we propose an intuitive approach inspired by Granovetter’s threshold model of social mobilisation (Granovetter 1978; Wiedermann et al. 2020).

In MaxParC, rather than stating their approval directly, voters again assign numerical ratings,  $0 \leq r_{ix} \leq 100$ . These are now interpreted as a ‘willingness to approve’, stating that “voter  $i$  approves option  $x$  if strictly less than  $r_{ix}$  percent of all voters do *not* approve  $x$ .” For example, a zero rating means “don’t approve no matter what”. A rating of 100 means “approve for sure”. Values in between mean “approve if enough others do so.” All ratings together result in a set of mutually dependent constraints that determine which voter ends up approving which options. It turns out that this set of equations can be solved quite easily using the methods from Granovetter (1978). For each option  $x$ , one first sorts the ballots in ascending order w.r.t. their rating of  $x$ . One then identifies the first ballot,  $i$ , that is preceded by strictly less than  $r_{ix}$  percent of all ballots. This ballot  $i$ , and all later ballots  $j$  (those with  $r_{jx} \geq r_{ix}$ ), “approve”  $x$ . In other words, the rating of this “pivot” voter  $i$  serves as a “cutoff” value for everyone’s approval of  $x$ . Graphically, the cutoff can easily be identified by finding the first intersection of the sorted ratings graph with its main diagonal, as in Fig. 3 (right).<sup>2</sup> After determining which voters approve which options, MaxParC then applies CUR on the so-calculated approval data. In case of remaining ties, the aggregated rating values decide.

<sup>2</sup> It would be harder to find a convincing solution if the conditions were not in terms of approval but in terms of where a voter’s *unique* vote goes—as in the “increasing group activity selection problems” of Darmann et al. (2012), Darmann (2019) and Abramowitz et al. (2021)—since then several maximal solutions might exist and monotonicity becomes an issue, see *Supplement 3.1.2*.





**Fig. 3** Group decision method MaxParC from the view of some voter Alice. Left box: Each rating value (thick colored needles coming in from the left) represents a conditional commitment by Alice to approve the respective option. Approval scores (nos. of voters approving an option) are represented by light bars coming in from the right, options are sorted by descending approval score. The rating of 13 for option C, for instance, is interpreted as saying that Alice approves C if and only if less than 13 percent of voters do *not* approve C. In other words, Alice is counted as approving an option if her rating needle overlaps with the approval score bar (green needles), and is otherwise counted as not approving the option (blue needles). Right diagrams: Approval scores can be determined graphically in a way similar to Granovetter (1978) by finding the leftmost intersection of the graph of ordered ratings with the main diagonal. Alice’s “vote” (= share of  $1/N$  of the winning probability) goes to the most-approved option approved by her (dark green needle). The resulting overall winning probabilities are shown as pie charts to the very left (color figure online)

Figure 3 illustrates the MaxParC procedure, which can (at least in principle) be performed using pen and paper. It is easy to see that if at least one option is rated positively by everyone, then among all such options the one with the largest aggregate rating will win for sure. In that case, MaxParC is like Range Voting restricted to the set of universally approved options. Only if no option gets all positive ratings will chance really play a role. For voters who are risk-averse, the potential use of chance can act as an incentive to identify good compromises and rate them positively, as in Heitzig and Simmons (2012). Indeed, MaxParC supports partial and full consensus in the test case of Fig. 2: all voters will give  $Z$  a slightly positive rating, resulting in its certain selection; if  $Z$  is not available, the  $F_{1,2}$  voters will give  $Y$  a rating slightly above  $S_3$  and thus transfer their votes and the corresponding winning probability safely from  $X_{1,2}$  to  $Y$ . In both cases, no voter has an incentive to reduce their rating of  $Y$  or  $Z$  in order to reserve their vote for their favourite option, because that would cause all others’ votes to go back to their favourites as well since their condition for approval would no longer be met.

In the next section, we provide some theoretical analysis of the formal properties of the Nash Lottery and MaxParC, which verifies that these two methods perform well in

terms of various qualitative criteria. In the subsequent sections, we use an agent-based model to make quantitative comparisons among various methods and a virtual lab experiment to understand the perceptions and preferences of American respondents with respect to MaxParC as compared to Plurality Voting. We limit the latter study to MaxParC due to the desirability of MaxParC relative to the Nash Lottery along various dimensions as well as financial and sampling constraints.

### 3 Formal analysis

Formally, we study (*probabilistic voting*) methods  $M : B(C)^E \rightarrow \Delta(C)$ , where  $C = \{1, \dots, k\}$  is a *choice set* of  $k > 1$  options (aka alternatives, candidates), denoted  $x, y, \dots$ ;  $E = \{1, \dots, N\}$  is an *electorate* of  $N > 2$  voters, denoted  $i, j, \dots$ ;  $B(C) \neq \emptyset$  is the set of possible (filled-in) *ballots*  $b$ ;  $B(C)^E$  is the set of (*ballot*) *profiles*  $\beta = (\beta_1, \dots, \beta_N)$ , specifying a ballot for each voter; and  $\Delta(C)$  is the set of possible (*winning*) *lotteries*  $\ell$ , i.e., probability distributions on  $C$ . We interpret  $M$  as a mechanism where the voters supply ballots  $\beta_i$  and then the winning option is drawn at random from the distribution  $M(\beta)$ . If  $B(C) = [0, 1]^C$ , so that  $b \in B(C)$  assigns ratings  $0 \leq b_x \leq 1$  to all options, we call  $M$  *ratings-based*. For convenience, we use abbreviations such as  $\sum_x$  and  $\sum_i$  for  $\sum_{x \in C}$  and  $\sum_{i \in E}$  where no confusion can arise.

The *Nash Lottery (NL)* is the ratings-based method  $M_{NL}$  defined as follows. Given  $\beta \in B(C)^E$ , for all  $m = 1, 2, 3, \dots$ , put  $r_i^m(\ell) = \sum_x \ell_x \sqrt[m]{\beta_{ix}}$  for all  $i$  and  $\ell \in \Delta(C)$ , and define a continuous, continuously differentiable, and weakly concave function  $S^m : \Delta(C) \rightarrow \mathbb{R}$  by  $S^m(\ell) = \sum_i \ln r_i^m(\ell) \in [-\infty, \infty)$ . The “Nash sum”  $S^1$  has a global maximum that is attained on a compact convex set  $T^1 = \arg \max_{\ell} S^1(\ell) \subseteq \Delta(C)$ . If  $T^1 = \{\ell\}$ , we put  $M_{NL}(\beta) = \ell$ . Otherwise we break the tie as follows: For  $m \geq 2$ ,  $S^m$  restricted to  $T^{m-1}$  has a global maximum that is attained on a compact convex set  $T^m \subseteq T^{m-1}$ . Then also  $T = \bigcap_{m=1}^{\infty} T^m$  is non-empty compact convex, hence Lebesgue-measurable, and hence has a well-defined unique centre of mass  $\ell^* = \int_T \ell d\ell / \int_T d\ell$  with  $\ell \in T$  because of the convexity. We then put  $M_{NL}(\beta) = \ell^*$ .

*Maximal Partial Consensus (MaxParC)* is the ratings-based method  $M_{MaxParC}$  defined as follows. Given  $\beta \in B(C)^E$ , for all  $x$ , let  $A_x$  (the set of voters who *approve*  $x$ ) be the largest subset  $A \subseteq E$  such that  $\forall i \in A : |A|/N + \beta_{ix} > 1$ . Let  $s_x = |A_x| + \sum_i \beta_{ix}/N$  (the *score* of  $x$ , basically the approval score with ties broken by average ratings). For all  $i$ , put  $A_i = \{x : i \in A_x\}$  (the set of options approved by  $i$ ) and  $W_i = \arg \max_{x \in A_i} s_x$  (the highest-scoring options approved by  $i$ , typically a singleton). Also, put  $A_{\emptyset} = E - \bigcup_x A_x$  (the set of *abstaining* voters). If  $A_{\emptyset} \neq E$ , put  $M_{MaxParC}(\beta)_x = \sum_{i: x \in W_i} 1/|W_i|(N - |A_{\emptyset}|)$ . Otherwise, let  $W = \arg \max_x s_x$  and put  $M_{MaxParC}(\beta)_x = 1_{x \in W}/|W|$  (if no-one approves anything, use a uniform distribution on the highest-scoring options, typically a unique option). Note that we have rescaled the ratings from  $[0, 100]$  to  $[0, 1]$  here (and in the *Supplement*) for convenience.

The method *Random Ballot* is defined by  $B(C) = C$  and  $M_{RB}(\beta)_x = |\{i : \beta_i = x\}|/N$ . *Full Consensus/Random Ballot (FC)* (“Voting method 1” from Heitzig and Simmons 2012) is defined by  $B(C) = C^2$ ,  $\beta_i = (\beta_i^1, \beta_i^2)$ ,  $M_{FC}(\beta)_x = 1$  if  $\forall i : \beta_i^1 = x$  (full consensus), and  $M(\beta)_x = |\{i : \beta_i^2 = x\}|/N$  if  $\nexists y \forall i : \beta_i^1 = y$  (use RB as

fallback if no full consensus exists). *Full Consensus/Random Ballot/Ratings (RFC)* (“Voting method 2” from Heitzig and Simmons 2012) is defined by  $B(C) = C^2 \times [0, 1]$  and  $M_{RFC}(\beta)_x = \frac{|A_x|}{N} + (1 - \frac{|A|}{N}) \frac{p_x}{N}$  (draw a voter, elect the proposed  $x$  if  $x$  is a potential full consensus, otherwise use the benchmark lottery). In this,  $A_x = \{i \mid \beta_i^1 = x \wedge \forall j : \beta_{jx}^3 \geq r_j\}$  (voters who propose  $x$  where  $x$  is a potential full consensus),  $r_j = \sum_x p_x \beta_{jx}^3 / N$  (effective ratings of the benchmark lottery),  $p_x = \{i : \beta_i^2 = x\}$  (benchmark lottery), and  $A = \bigcup_x A_x$  (voters who propose a potential full consensus). *Supplement 2.2* has more details on these methods and the formal framework.

### 3.1 Consistency, manipulability and fairness properties

We restrict ourselves to a subjectively chosen set of basic properties here that can be proved or disproved easily, leaving a more comprehensive axiomatic treatment to future work.

Let  $\Pi(A)$  be the set of all permutations on  $A$ . We say that a method  $M$  is *anonymous* iff for all  $\beta \in B(C)^E$  and  $\sigma \in \Pi(E)$ ,  $M(\beta \circ \sigma) = M(\beta)$ . We call  $M$  *neutral* iff for all  $\beta$  and  $\sigma \in \Pi(C)$ ,  $M(\beta') = M(\beta) \circ \sigma$ , where  $\beta'_i = \beta_i \circ \sigma$  for all  $i$ . We call a ratings-based  $M$  *Pareto-efficient (w.r.t. stated preferences)* iff  $M(\beta)_y = 0$  whenever  $\exists x \forall i : \beta_{iy} < \beta_{ix}$ . We call a ratings-based  $M$  *almost surely resolute* iff the set  $\{\beta \in B(C)^E : |\{x : M(\beta)_x > 0\}| > 1\}$  has Lebesgue-measure zero.

We consider two versions of ‘mono-raise’ monotonicity from Woodall (1997). We call a ratings-based  $M$  *weakly mono-raise monotonic* iff  $M(\beta)_x \geq M(\beta')_x$  whenever  $\exists i : \beta_{ix} > \beta'_{ix} \wedge (\forall y \neq x : \beta_{iy} = \beta'_{iy}) \wedge \forall j \neq i : \beta_j = \beta'_j$  (decreasing an option’s rating does not increase its chances). We call a ratings-based  $M$  *strongly mono-raise monotonic* iff  $M(\beta)_y \leq M(\beta')_y$  whenever  $\exists i \exists x \neq y : \beta_{ix} > \beta'_{ix} \wedge (\forall z \neq x : \beta_{iz} = \beta'_{iz}) \wedge \forall j \neq i : \beta_j = \beta'_j$  (decreasing an option  $x$ ’s rating does not decrease any other option  $y$ ’s chances). Note that the strong version implies the weak one.

Regarding manipulability by introducing or removing options, we only study two properties. We call a neutral, ratings-based  $M$  *option-removal monotonic* iff  $\forall y < k : M(\beta)_y \leq M(\beta')_y$  whenever  $\beta \in B([k])^E$  and  $\beta' \in B([k - 1])^E$  with  $\forall i : \beta'_i = \beta_{i|[k-1]}$  (removing an option does not decrease any other option’s chances). We call a neutral, ratings-based  $M$  *independent from losing options* iff  $\forall y < k : M(\beta)_y = M(\beta')_y$  whenever  $\beta \in B([k])^E$  and  $\beta' \in B([k - 1])^E$  with  $\forall i : \beta'_i = \beta_{i|[k-1]}$  and  $M(\beta)_k = 0$  (removing an option with no chance does not change any other option’s chances). Note that option-removal monotonicity implies independence from losing options.

Generalizing the ‘Core Fair Share’ axiom from Aziz et al. (2019) to our framework, we say an anonymous method  $M$  *allocates power proportionally* (aka *is proportional*) iff  $\forall x \forall m \leq N \exists \beta^m \in B(C)^{[m]} \forall \beta^{N-m} \in B(C)^{[N] \setminus [m]} : M(\beta^m, \beta^{N-m})_x \geq m/N$  (a coalition of  $m$  voters have a way of voting that guarantees option  $x$  a chance at least  $m/N$ ). In contrast, we call  $M$  *majoritarian* iff  $\forall x \forall m > N/2 \exists \beta^m \in B(C)^{[m]} \forall \beta^{N-m} \in B(C)^{[N] \setminus [m]} : M(\beta^m, \beta^{N-m})_x = 1$  (a strict majority can guarantee  $x$  a sure win). Obviously, a majoritarian method (blue line in Fig. 1) cannot allocate power proportionally (green line in Fig. 1).

**Proposition 1**  $M_{NL}$ ,  $M_{MaxParC}$ ,  $M_{RB}$ ,  $M_{FC}$ , and  $M_{RFC}$  are anonymous, neutral, and allocate power proportionally.  $M_{NL}$  and  $M_{MaxParC}$  are Pareto-efficient.  $M_{MaxParC}$  is almost surely resolute.  $M_{MaxParC}$  is strongly mono-raise monotonic and option-removal monotonic, but  $M_{NL}$  is neither.

Plurality Voting, Approval Voting, Range Voting, Instant-Runoff Voting, and all Condorcet-consistent methods are majoritarian, independently of which tie- or cycle-breaking method is used.

**Proof** Anonymity and neutrality are straightforward from the definitions.

Pareto-efficiency of  $M_{NL}$ : We show that for lotteries where  $\ell_y > 0$ ,  $S^1(\ell)$  can be increased by moving  $y$ 's probability mass to  $x$ . Let  $\ell = M_{NL}(\beta)$ , assume  $\ell_y > 0$ , and define  $\ell' \in \Delta(C)$  by  $\ell'_x = \ell_x + \ell_y > \ell_x$ ,  $\ell'_y = 0 < \ell_y$ , and  $\ell'_z = \ell_z$  for all other  $z$ . Then  $\forall i : r_i^1(\ell') > r_i^1(\ell) \geq 0$ , hence  $S^1(\ell') > S^1(\ell)$  and  $\ell \notin T^1 \supseteq T \ni M_{NL}(\beta)$ , so  $M_{NL}(\beta) \neq \ell$  after all.

Pareto-efficiency of  $M_{MaxParC}$ : We show that  $y$  is no-one's highest-scoring approved option. We have  $A_y \subseteq A_x$ , thus  $a'_y < a'_x$ , hence  $y \notin W$  and  $\forall i : y \notin A'_i$ , so  $M_{MaxParC}(\beta)_y = 0$ .

$M_{NL}$  is proportional: Assume  $\beta_{ix}^m = 1, \forall y \neq x : \beta_{iy}^m = 0, \ell^* = M_{NL}(\beta)$ , and  $\ell_x^* < m/N$ . We show that moving an infinitesimal probability mass towards  $x$  will strictly increase  $S^1(\ell)$ , so that  $\ell^* \notin T^1$ , a contradiction. For  $\epsilon \geq 0$ , let  $\ell_x^\epsilon = (1 - \epsilon)\ell_x^* + \epsilon$  and  $\ell_y^\epsilon = (1 - \epsilon)\ell_y^*$  for all  $y \neq x$ . Then indeed

$$\begin{aligned} \partial_\epsilon S^1(\ell^\epsilon)|_{\epsilon=0} &= \sum_i \frac{\sum_z \partial_\epsilon \ell_z^\epsilon|_{\epsilon=0} \beta_{iz}}{\sum_z \ell_z^* \beta_{iz}} = \sum_i \frac{(1 - \ell_x^*)\beta_{ix} - \sum_{y \neq x} \ell_y^* \beta_{iy}}{\ell_x^* \beta_{ix} + \sum_{y \neq x} \ell_y^* \beta_{iy}} \\ &> \sum_{i \leq m} \frac{(1 - m/N)}{m/N} + \sum_{i > m} \frac{(1 - m/N) \times 0 - \sum_{y \neq x} \ell_y^* \beta_{iy}}{m/N \times 0 + \sum_{y \neq x} \ell_y^* \beta_{iy}} = 0. \end{aligned}$$

$M_{MaxParC}$  is proportional: With the same  $\beta^m$ , we get  $\forall i \leq m : W_i = A_i = \{x\}$ , hence  $M_{MaxParC}(\beta)_x \geq m/N$ .

$M_{RB}, M_{FC}, M_{RFC}$  are proportional: It is easy to see that  $\beta_i = x$  in  $M_{RB}$ ,  $\beta_i^1 = \beta_i^2 = x$  in  $M_{FC}$ , and  $\beta_i^1 = \beta_i^2 = x, \beta_i^3 = 1 \wedge \forall y \neq x : \beta_i^3 = 0$  in  $M_{RFC}$  suffice to make  $M(\beta)_x \geq m/N$ .

Almost sure resoluteness of  $M_{MaxParC}$ : Except for a set of measure zero, we have  $\forall i, x : \beta_{ix} > 0$  and  $\forall x \neq y : s_x \neq s_y$ , which implies  $\exists x^* \forall y \neq x^* : s_{x^*} > s_y$ ,  $\forall i : A_i = C, W_i = \{x^*\}$ , and hence  $M_{MaxParC}(\beta)_{x^*} = 1$ .

Strong mono-raise monotonicity of  $M_{MaxParC}$ : We show that when going from  $\beta'$  to  $\beta$ , only  $x$  can gain approval or score and thus winning probability. Let  $A', s', W'$  denote the versions of  $A, s, W$  that are based on  $\beta'$  instead of  $\beta$ . Then  $A_x \supseteq A'_x, s_x > s'_x, \forall y \neq x : A_y = A'_y \wedge s_y = s'_y, \forall j : A_j \in \{A'_j, A'_j \cup \{x\}\} \wedge W_j \in \{\{x\}, W'_j, W'_j \cup \{x\}\}, A_\emptyset \subseteq A'_\emptyset$ , and  $W \in \{\{x\}, W', W' \cup \{x\}\}$ . Hence,  $\sum_{i: y \in W'_i} 1/|W_i|(N - |A_\emptyset|) \leq \sum_{i: y \in W_i} 1/|W_i|(N - |A_\emptyset|)$  and  $M_{MaxParC}(\beta)_y = 1_{y \in W}/|W| \leq M_{MaxParC}(\beta)_y = 1_{y \in W'}/|W'|$  for all  $y \neq x$ .

Option-removal monotonicity of  $M_{MaxParC}$ : Similarly, we show that when concatenating the ratings for option  $k$  to  $\beta'$  to get  $\beta$ , no  $y \neq k$  can gain approval

or score and thus winning probability: Again,  $\forall y \neq k : A_y = A'_y \wedge s_y = s'_y$ ,  $\forall j : A_j \in \{A'_j, A'_j \cup \{k\}\}$ . The rest of the argument is the same as for strong mono-raise monotonicity with  $x = k$ .

Counterexample to  $M_{NL}$  being strongly mono-raise monotonic: Consider  $N = 2$ ,  $k = 3$ ,  $x = 1$ ,  $y = 3$ ,  $\beta_1 = (1, 1/6, 0)$ ,  $\beta_2 = (0, 3/4, 1)$ ,  $\beta'_1 = (1/2, 1/6, 0)$ , and  $\beta'_2 = \beta_2$ . Then  $M(\beta) = (1/2, 0, 1/2)$  and  $M(\beta') = (1/4, 3/4, 0)$ , hence  $M(\beta)_3 > M(\beta')_3$  although only a rating for option 1 was increased.

Counterexample to  $M_{NL}$  being option-removal monotonic: Consider the same  $\beta$  and now  $y = 1$ . Then  $M(\beta') = (2/5, 3/5)$ , hence  $M(\beta)_1 > M(\beta')_1$ .

Other options being majoritarian: It is straightforward to see that for all of those methods, if all  $i \in [m]$  specify  $x$  as the unique top-ranked option on  $\beta_i$ , or put  $\beta_{ix} = 1$  and  $\forall y \neq x : \beta_{iy} = 0$  in case of Range Voting, then  $x$  will win for sure because  $m > N/2$ . □

In *Supplement 3.1.4* we also show that replacing the logarithm by some other function in the definition of  $M_{NL}$  will destroy proportionality. Extensive numerical simulations suggest further:

**Conjecture 1**  $M_{NL}$  is weakly mono-raise monotonic, independent from losing options,<sup>3</sup> but not almost surely resolute.

A complete formal treatment of *clone-proofness* is beyond the scope of this paper. We only discuss the case here where for profile  $\beta$ , there are no relevant ties and then an additional option  $k + 1$  is introduced whose ratings are exactly the same as for some existing option, say  $k$ , giving a profile  $\beta'$ . More precisely, if for  $\beta$ , the tie-breaking set  $T$  of  $M_{NL}$  is a singleton  $\{\ell^*\}$ , then with  $\beta'$ , we have  $T = \{\ell' \in \Delta([k + 1]) : \ell'|_{[k-1]} = \ell^*|_{[k-1]}\}$  and thus  $M_{NL}(\beta')|_{[k-1]} = M_{NL}(\beta)|_{[k-1]}$ . Similarly, if for  $\beta$ , all scores  $s_x$  are distinct, then with  $\beta'$ , all  $W_i \in \{\{x\} : x < k\} \cup \{\{k, k + 1\}\}$  and thus again  $M_{MaxParC}(\beta')|_{[k-1]} = M_{MaxParC}(\beta)|_{[k-1]}$ . In other words, adding an exact clone does not change the chances of non-cloned options unless there are already too many ties.

### 3.2 Consensus-supporting properties

We restrict our formal analysis to the case presented in Fig. 2 of the introduction, treating it as a one-shot, simultaneous-move, “factional” game in which the players are not the individual voters but the factions  $F_1, F_2, F_3$ . Without loss of generality, assume  $C = [5] = \{X_1, X_2, X_3, Y, Z\}$ ,  $E = F_1 \cup F_2 \cup F_3$ ,  $F_1 = \{1, \dots, N_1\}$ ,  $F_2 = \{N_1 + 1, \dots, N_1 + N_2\}$ ,  $F_3 = \{N_1 + N_2 + 1, \dots, N\}$ , and  $N_1 \geq N_2$ , where  $N_t = |F_t|$ ,  $N = N_1 + N_2 + N_3$ , and  $n_t = N_t/N$  for  $t = 1 \dots 3$ . The (pure) strategy  $s_t$  of faction  $F_t$  is then to choose the ballots of all its members,  $s_t \in B(C)^{F_t}$ , resulting in profile  $\beta = (s_1, s_2, s_3) \in B(C)^E$ . Player  $F_t$ 's payoff for lottery  $\ell \in \Delta(C)$ ,  $\pi_t(\ell)$ , is the sum of its members' expected utility functions,  $\sum_{i \in F_t} \sum_x \ell_x u_{ix}$ . Their expected payoff resulting from profile  $\beta$  under method  $M$  is then  $\pi_t(M(\beta)) = \sum_x M(\beta)_x \sum_{i \in F_t} u_{ix}$ .

Assume that individual utilities are as follows:  $\forall i \in F_t, t' \neq t : u_{iX_t} = 1 \wedge u_{i(X_{t'})} = 0 \wedge u_{iZ} = v_Z < 1$ ,  $\forall i \in F_1 \cup F_2 : u_{iY} = v_Y < 1$ , and  $\forall i \in F_3 : u_{iY} = 0$ .

<sup>3</sup> The case where  $T^1$  is a singleton is readily proved, only the tie-breaking case is the hard part.

Finally, assume  $v_Z > \max\{n_1, n_2, n_3\}$ , and  $v_Z > (n_1 + n_2)v_Y$ . In other words, all factions prefer  $Z$  to the benchmark lottery  $\ell_{X_i}^b = n_i, \ell_Y^b = \ell_Z^b = 0$ ; and factions  $F_1, F_2$  prefer  $Z$  also to the “partial consensus” lottery  $\ell_{X_3}^c = n_3, \ell_Y^c = n_1 + n_2$ .

We now show that under MaxParC, the game has strong equilibria in which full or partial consensus is realized.

**Proposition 2** *In the above situation, if  $M = M_{MaxParC}$  and  $(n_1 + n_2)v_Y \leq v_Z$ , then the factional game has a strong Nash equilibrium  $\beta$  in which  $M(\beta)_Z = 1$ .*

**Proof** First, we note that under any possible resulting lottery  $\ell$ , total payoffs would be  $\Pi(\ell) = N_1\ell_{X_1} + N_2\ell_{X_2} + N_3\ell_{X_3} + Nv_Z\ell_Z + (N_1 + N_2)v_Y\ell_Y \leq Nv_Z$ , hence the grand coalition of all three factions has no way to improve their payoff if  $M(\beta)_Z = 1$ .

We show that the following profile fulfils the claim:  $\forall i \in F_i, t' \neq m : \beta_{iX_i} = 1 \wedge \beta_{iZ} = 1/N \wedge \beta_i(X_{t'}) = \beta_{iY} = 0$ . With this  $\beta$ , factional and total payoffs are  $\pi_i(M(\beta)) = N_i v_Z$  and  $\Pi(M(\beta)) = N v_Z$ .

Faction  $F_3$  alone could increase their payoff  $\pi_3(\ell) = N_3(\ell_{X_3} + v_Z\ell_Z)$  only by moving probability mass from  $Z$  to  $X_3$ . To do so, they need to put  $\beta_{iZ} = 0$  for some  $i \in F_3$ , but this would result in  $A_Z \cap (F_1 \cup F_2) = \emptyset$  because of  $\beta_{jZ} = 1/N$  for all  $j \in F_1 \cup F_2$ . The resulting  $\ell$  would have  $\ell_{X_1} \geq n_1, \ell_{X_2} \geq n_2$ , and  $\ell_{X_3} + \ell_Z \leq n_3$ , hence  $\pi_3(\ell) \leq N_3 n_3 < N_3 v_Z$ , which is not an improvement. Hence  $F_3$  has no profitable deviation.

Similarly, faction  $F_1$  alone could increase  $\pi_1(\ell) = N_1(\ell_{X_1} + v_Z\ell_Z + v_Y\ell_Y)$  only by moving mass from  $Z$  to either  $X_1$  or  $Y$ . Again, this would require  $\beta_{iZ} = 0$  for some  $i \in F_1$ , resulting in  $A_Z \cap (F_2 \cup F_3) = \emptyset$  and thus  $\ell_{X_2} \geq n_2, \ell_{X_3} \geq n_3, \ell_{X_1} + \ell_Z + \ell_Y \leq n_1$ , and  $\pi_1(\ell) \leq N_1 n_1 < N_1 v_Z$ , not an improvement. So  $F_1$  has no profitable deviation. The same holds for  $F_2$  by symmetry.

The coalition  $F_1 \cup F_2$  could increase  $\pi_{12}(\ell) := \pi_1(\ell) + \pi_2(\ell) = N_1\ell_{X_1} + N_2\ell_{X_2} + (N_1 + N_2)(v_Z\ell_Z + v_Y\ell_Y)$  above its value of  $(N_1 + N_2)v_Z$  only by moving mass from  $Z$  to either  $X_1, X_2$ , or  $Y$ . This requires  $\beta_{iZ} = 0$  for some  $i \in F_1 \cup F_2$ , hence  $A_Z \cap F_3 = \emptyset, \ell_{X_3} \geq n_3$ , and  $\ell_{X_1} + \ell_{X_2} + \ell_Z + \ell_Y \leq n_1 + n_2$ . Under these constraints,  $\max_{\ell} \pi_{12}(\ell) < (N_1 + N_2)v_Z$  because of the premise  $(n_1 + n_2)v_Y \leq v_Z$ .

Similarly, coalition  $F_1 \cup F_3$  could increase  $\pi_{13}(\ell) := \pi_1(\ell) + \pi_3(\ell) = N_1(\ell_{X_1} + v_Y\ell_Y) + N_3\ell_{X_3} + (N_1 + N_3)v_Z\ell_Z$  above its value of  $(N_1 + N_3)v_Z$  only by moving mass from  $Z$  to either  $X_1, X_3$ , or  $Y$ . This requires  $\beta_{iZ} = 0$  for some  $i \in F_1 \cup F_3$ , hence  $A_Z \cap F_2 = \emptyset, \ell_{X_2} \geq n_2$ , and  $\ell_{X_1} + \ell_{X_3} + \ell_Z + \ell_Y \leq n_1 + n_3$ . Under these constraints,  $\max_{\ell} \pi_{13}(\ell) < (N_1 + N_3)v_Z$  because of the premise  $n_1 v_Y \leq v_Z$ . The case for coalition  $F_2 \cup F_3$  is completely analogous by symmetry.

Hence no coalition has an improving deviation. □

Note that this holds even if  $n_1 > 1/2$ , in which case all majoritarian methods would give  $X_1$  for sure in equilibrium because  $F_1$  strictly prefers this to any other lottery and can enforce it.

**Proposition 3** *If  $M = M_{MaxParC}$ ,  $v_Y > \max\{N_1, N_2\}/(N_1 + N_2)$ , and option  $Z$  is removed, then the factional game has a strong Nash equilibrium  $\beta$  in which  $M(\beta)_Y = n_1 + n_2$  and  $M(\beta)_{X_3} = n_3$ .*

**Proof** Completely analogous to the previous proof, only that this time  $\forall i \in F_1 \cup F_2 : \beta_{iY} = (N_3 + 1)/N$ . □

Because  $M_{NL}$  has additional complexity due to the involved nonlinear optimization and tie-breaking procedure, we do not prove the corresponding claims for  $M_{NL}$  but only conjecture them and give some supporting argument. Based on the results from Heitzig and Simmons (2012), we also make a conjecture about  $M_{FC}$  and  $M_{RFC}$ :

**Conjecture 2** *In the above situation, if  $M \in \{M_{NL}, M_{FC}, M_{RFC}\}$  and  $(n_1 + n_2)v_Y \leq v_Z$ , then the factional game has a strong Nash equilibrium  $\beta$  in which  $M(\beta)_Z = 1$ . Also, if option Z is removed and  $v_Y > \max\{N_1, N_2\}/(N_1 + N_2)$ , then the factional game has a strong Nash equilibrium  $\beta$  in which  $M_{NL}(\beta)_Y = n_1 + n_2$  and  $M_{NL}(\beta)_{X_3} = n_3$ , while for  $M_{FC}$  and  $M_{RFC}$  no such strong Nash equilibrium exists.*

Indeed, for the first part and  $M = M_{NL}$ , let us put  $\beta$  as above with the exception that  $\forall i \in F_t : \beta_{iZ} = n_t$ . Let  $\ell = M(\beta)$  and  $\ell_Z^Z = 1$ , so that  $S^1(\ell^Z) = \sum_t N_t \ln n_t \geq \sum_t N_t \ln n_t$  and  $S^2(\ell^Z) = \sum_t N_t \ln \sqrt{n_t} > \sum_t N_t \ln n_t$ . To see that  $\ell = \ell^Z$ , i.e., Z will win for sure, we note that moving an infinitesimal amount from  $\ell^Z$  towards an  $\ell' \neq \ell^Z$  with  $\ell'_Z = 0$  would cause a weak decline in  $S^1$  and a strict one in  $S^2$ , so that the resulting  $\ell \notin T$ . This is because

$$\begin{aligned} \partial_\epsilon \sum_t N_t \ln(\epsilon \ell'_{X_t} + (1 - \epsilon)n_t)|_{\epsilon=0} &= \sum_t N_t \frac{\ell'_{X_t} - n_t}{n_t} \\ &\leq \sum_t N \ell'_{X_t} - N \leq N - N = 0, \end{aligned}$$

and analogously for  $S^2$ , with  $<$  instead of  $\leq$ . For the second part, we put  $\forall i \in F_t : \beta_{iY} = n_t/(n_1 + n_2)$  for  $t = 1, 2$ , which one can similarly show to lead to the claimed output. However, to prove that no deviation by a single faction or pair of factions pays off will be more difficult for  $M_{NL}$  than for  $M_{MaxParC}$  because we cannot rely on strong monotonicity and have to distinguish several cases depending on whether the optimization problem has an interior or boundary solution and whether a tie breaker  $S^m$  with  $m > 1$  is involved. *Supplement 3.1.5* has some additional calculations in this direction, and also gives an example where in the presence of several disjoint partial consensus options, the corresponding ballot profile can fail to be a strong equilibrium while still being coalition-proof.

### 4 Agent-based simulations

To assess the potential costs of achieving fairness and consensus in terms of welfare loss, voter satisfaction, and amount of randomization, we complement our theoretical analyses with empirical evidence from a large agent-based simulation, which can be understood as a complex, computer-assisted thought experiment.

**Experiment setup.** In a diverse sample of almost 3 million hypothetical group decision problems, we compared the performance of the Nash Lottery and MaxParC to that of



eight other methods. We chose five common deterministic majoritarian group decision methods: Plurality Voting (PV, aka ‘first past the post’), Approval Voting (AV), Range Voting (RV), Instant Runoff Voting (IRV, aka ‘ranked choice voting’), and the ‘Simpson–Kramer’ method (aka ‘Simple Condorcet’, SC). We also included three non-deterministic proportional methods: ‘Random Ballot’ (RB, aka ‘Lottery Voting’), ‘Full Consensus/Random Ballot’ (FC), and ‘Full Consensus/Random Ballot/Ratings’ (RFC). The latter two are precursors of MaxParC that only support full consensus but not partial consensus (Heitzig and Simmons 2012). To generate a diverse set of decision problems, we created random combinations of options, resulting in choice sets that varied in number and compromise potential, and groups of voters, which varied in size, individual preference distributions, and risk attitudes.

We used various preference models from behavioural economics and the spatial theory of voting. The ‘unif’ model is similar to the ‘impartial culture’ (Laslier 2010), which assumes that utility values are uniformly distributed on the interval  $[0, 1]$  with no structure whatsoever. In the block model (BM), voters form blocks with high internal preference correlation, with no correlation between blocks, so that utility is a sum of a block-specific normally distributed term and a smaller individual-specific normally distributed term. The spatial models LA, QA, GA are variants of those used in the spatial theory of voting (Carroll et al. 2013), where voters and options are represented as normally distributed points in a low-dimensional “policy space”, and utility depends on a voter’s distance from an option, either in a linear (LA), quadratic (QA) or Gaussian (GA) fashion.

For each combination of decision problem and group decision method, we simulated several opinion polls, a main voting round, and an interactive phase where ballots could be modified continuously for strategic reasons. We assumed various mixtures of behavioural types among our simulated voters: lazy voting, sincere voting, individual heuristics, trial and error, and coordinated strategic voting. For each decision problem, we computed measures of the social welfare, randomness, and voter satisfaction for all group decision methods, and identified which group decision methods different voter types would prefer (see *Supplement 2.3* for details).

**Results.** As can be expected from the definition of ‘majoritarian method’, some majority of simulated voters (namely the ones whose preferences were being enacted under the majoritarian method) preferred the results of the majoritarian methods over those of the proportional ones. Within the group of proportional methods, voters preferred MaxParC over the other methods on average, which might be related to its lower amount of randomness and higher welfare (see below). Among the majoritarian methods, there was no clear preference over the methods. Individual voters’ satisfaction—normalized to zero for their least-preferred option and to unity for their favourite—averaged around 67% for PV, AV, RV, and IRV; 61% for SC, NL, and MaxParC; and still 57% for RB, FC, and RFC. So MaxParC achieved 91% of the level of voter satisfaction achieved by the best deterministic methods, and about the same as typical Condorcet methods. This indicates that MaxParC was indeed able to incentivise simulated voters to support good compromise options.

Regarding randomness, the outcome lotteries under MaxParC had only about 56% of the Shannon entropy of those under RB, while NL produced about 77% of RB’s

entropy. For Rényi entropy, this was similar. In MaxParC, the largest winning probability was about 69% on average, in NL it was only about 56%. In particular, MaxParC led to a deterministic winner in about 19% of cases, NL only in about 11% of cases. These findings show that MaxParC used significantly less randomization than NL, which can be explained by its higher concentration of winning probability on the most-approved option (which receives its full approval score as winning probability).

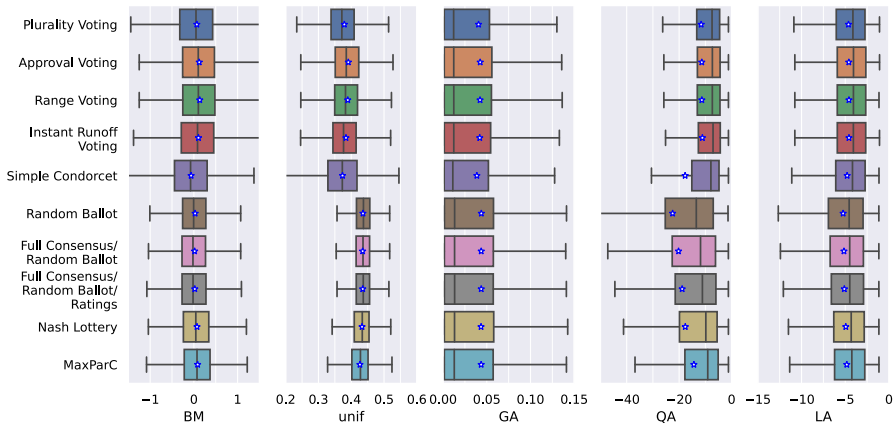
The deterministic methods produced somewhat higher welfare on average, but for some preference models and welfare metrics, the non-deterministic methods matched or outperformed the deterministic methods. In two of the preference models ('BM' and 'unif'), the majoritarian methods generated slightly larger absolute utilitarian welfare and smaller absolute egalitarian welfare values than the proportional methods. On the intermediate Gini–Sen welfare metric, the proportional methods beat the majoritarian ones in the 'unif' preference model, but were beaten by the majoritarian methods in two other preference models ('QA' and 'LA'), see Fig. 4. This confirms the intuition that the more inequality-averse the welfare metric, the better non-deterministic methods will perform because they can equalize individual expected utility by suitable randomization.

The striking discrepancy between the 'unif' and 'QA' preference models is due to their extreme difference in structure. In the 'unif' model, the benchmark lottery, and the lotteries resulting from proportional methods, are evaluated by all voters similarly, at around  $1/2$  utility and with a variance that decreases with larger numbers of voters and options. This leads to welfare values relatively close to 0.5 for the proportional methods. However, a voter's utility for any given option varies from 0 to 1, leading to Gini–Sen welfare of only about  $1/3$  for the mostly deterministic majoritarian methods, due to this inequality. In the 'QA' spatial model, on the other hand, utility functions are strictly convex functions of options' positions in policy space, so if an option is spatially central, it is typically strictly preferred by most voters to an option lottery. At the same time, spatial models often have central options that are Condorcet winners (especially if the dimension is small and there are many options) and are thus likely selected by the majoritarian methods. Hence the winners of majoritarian methods tend to be preferred by most voters to option lotteries, in particular to those resulting from proportional methods, in spatial models with convex utilities like the 'QA' model.

When fitting regression models to estimate the welfare metrics, we found that, in general and across all group decision methods, both a larger dimension of the policy space in the spatial theory of voting, and a greater level of spatial voter heterogeneity resulted in decreased absolute welfare. As expected, having more options increased welfare on average.

For most of the simulation results, preference distributions had a larger effect than behavioural type or the amount of interaction (number of polling rounds and presence of interactive phase). Surprisingly, for none of the decision methods, strategic voters gained a clear advantage over lazy voters, and risk attitudes played a minor role. See *Supplement 3.2* for more detailed results.

Finally, for each decision problem we also compared (i) the difference between individual voters' utilities when a particular group decision method was used, and (ii) the difference between voters' average utility under different group decision methods. We found that in more than 75% of all decision problems, the utility difference between



**Fig. 4** Simulated welfare effects of various group decision methods. Distribution of final Gini–Sen welfare across 2.5 mio. agent-based simulations by group decision method (rows), for five different models of how voter preferences might be distributed (columns). See *Supplement* for definitions and detailed additional results

the average and worst-off voter when using Range Voting—the best deterministic method considered—was at least seven times as large as the difference in average voter utility between the results of Range Voting and MaxParC. One interpretation of this findings is that the welfare costs of fairness and consensus are small compared to the inequality costs of majoritarianism.

Inspired by the latter finding, we propose to study more generally an indicator of the “inequality-scaled utilitarian costs” of using a method  $M$  for a decision problem with utility functions  $u_i$  and getting a lottery  $\ell$ , in spirit similar to ratios used by Merrill (1988). It would divide the shortfall in utilitarian welfare from using  $M$  by a measure of the inequality that would result from the utilitarian solution  $x^*$ ,

$$\phi(\ell) = \frac{\sum_i u_{ix^*} - \sum_i \sum_x \ell_x u_{ix}}{\sum_i u_{ix^*} - N \min_i u_{ix^*}}, \quad \text{where } x^* = \arg \max_x \sum_i u_{ix}.$$

## 5 Virtual lab experiment

Here, we briefly describe the sample, methods and results of a virtual lab experiment aimed at assessing individual-level predictors of choice between two different voting methods—Plurality Voting and MaxParC—in an incentivized setting with material consequences. The study was performed in 12 data collection waves, run between March 19 and April 12, 2021 and distributed across weekdays and times of day. The study proceeded as follows: 1. Questionnaire; 2. Instructions and 8 rounds of group decisions with Plurality Voting; 3. Instructions and 8 rounds of group decisions with MaxParC; 4. Voting method free choice (preference elicitation) followed by 2 rounds of play; and 5. Debriefing questionnaire. Hypotheses, experimental design, sampling

and analysis plan were pre-registered (<https://osf.io/2568f/>). Please refer below and to the *Supplement* for additional details.

**Sample.** We ran a 30-min experiment online using the Dallinger platform for online behavioural experiments<sup>4</sup> with participants from Amazon Mechanical Turk (MTurk). 616 participants completed the experiment, and 1096 individuals started the experiment. Of these, 420 did not pass the attention check and 60 did not complete the experiment, resulting in 616 experimental participants. 303 of these participants were randomly assigned to the ‘high inequality’ treatment, and 313 to the ‘low inequality’ treatment. 42.7% of our sample is female, and the average age is 37.4. The sample size was determined by a power calculation, pragmatic constraints, and an expected attrition rate of 50%. We restricted our sample to participants who were at least 18 years of age, English speakers, and had an MTurk approval rating of at least 95. We excluded participants who failed the comprehension questions following the instructions from participation in subsequent rounds, and we dropped all participants who did not complete both voting methods and the preference elicitation rounds from all analyses. Participants received a base payment of \$0.10 for participating, compensation of \$2 upon submission of the task, and a performance-dependent bonus payment. The average and maximum possible total payments were \$8.58 and \$13.10, respectively. This study was approved by the Princeton University Institutional Review Board.

**Voting Game.** Participants, playing in groups, faced five different options for how to distribute points or earnings among members of the group. The study included a single between-subjects group-level treatment with two treatment arms: high and low distributive inequality conditions (see Tables S1 and S2 in the Supplementary Materials). The experiment was conducted across 12 waves, with each wave corresponding to one of the two treatments. Waves were scheduled to ensure a similar distribution of weekdays and experiment times across the two treatment arms. Participants were unaware of the treatment condition and were assigned to treatment based on the wave they participated in. Once in a treatment group, they were assigned a random “advantage value”, which was drawn uniformly between zero and one and held constant for the length of play. The relative ordering of the advantage values in a group determined how likely a participant was to be favored by the different payoff distribution options they faced in the voting game (see Supplementary Materials for additional information). E.g., the higher the advantage value, the more often a subject would find that the majority option in a certain round would be profitable for them.

During the voting rounds, participants made a series of 18 group decisions: 8 with Plurality Voting, 8 with MaxParC, and 2 with their choice of voting method, either PV or MaxParC. They were randomly matched with four other participants in each round, forming temporary groups of five participants whose composition changed across rounds. Participants were anonymous, unable to communicate and unlikely to encounter the same group twice. They were given full information about the group formation protocols (see Supplementary Materials for additional information. In order for the game to continue smoothly despite differences in response times and potential dropouts, we simulated participants (“bot players”) when necessary to maintain groups

<sup>4</sup> <https://github.com/Dallinger/Dallinger>

of five, which was made transparent to the subjects at the start of those rounds. Simulated participants made rational, self-interested decisions by playing the focal Nash equilibrium strategy. Approximately 46% of rounds using PV as the voting method and 53% of rounds using MaxParC included at least one simulated participant.

In each round, the five participants were presented with five options from which they had to collectively choose one that would be implemented for pay. Each option represented a distribution of points among the five participants in the group for that round. Each point earned in any round led to a payout of USD 0.01 at the end of the game. Payoffs thus depended on advantage values of the five participants, how an individual filled their ballot, how others in their group filled their ballots, and how the voting method processed the ballots.

The payoffs in each option were designed so that in a group of purely self-interested participants, three of the five (the “majority” faction) would favor a “majority option”, while the other two would favor one of two “minority options”. By design, no self-interested participant should favor the “partial” or “full compromise” options that were offered. In contrast, participants who are not purely self-interested but are sufficiently inequality-averse might favor the “compromise” options over the “majority” option since they lead to more equal payoff distributions. A participant’s advantage value determined the likelihood that they belonged to the “majority faction”. In addition, in the high inequality treatment, also the payoff amounts were positively correlated with the participants’ advantage values (see *Supplement 2.4.1* for additional details about advantage values, payoffs, and choice options).

Under Plurality Voting, the “majority faction” is expected to vote for the clearly discernable “majority option” if they are self-interested voters. By contrast, under MaxParC all self-interested voters should give the “full compromise option” a positive rating, resulting in its sure election and thus a more equal payoff distribution.

**Hypotheses and analyses.** Our primary analyses and hypotheses focus on the revealed preference over voting methods in the preference elicitation portion of the experiment. Our primary dependent variable for our individual analyses is the choice of MaxParC or PV in the “free choice” portion of the game, and the aggregate analyses use the share  $S$  of participants selecting MaxParC in these rounds. A game-theoretic analysis suggests that if all participants were self-interested and rational, 44.8% would select MaxParC over PV under the given payoff and advantage value structure. However, behavioral economists and psychologists have found that individuals often make choices that are not purely self-interested, and also tend to avoid risks and can be cognitively miserly, avoiding complex tasks where possible (Fehr and Schmidt 2006; Falk et al. 2018). As voting rationally under MaxParC might appear more complex than under Plurality when given only short time, this might affect subjects’ method choice. Additionally, political scientists and decision theorists sometimes assume additional costs associated with methods that are too complex (Fishburn 1974) or involve considerable amounts of chance (Gill and Gainsou 2002).

These observations resulted in three pre-registered hypotheses:

*Hypothesis 1a. Mild Complexity or Chance Aversion.*  $S$  will be less than the self-interested rational fraction of 44.8%, collapsing across treatment assignment.

*Hypothesis 1b. Complexity or Chance Aversion.* More risk-averse individuals or those who found MaxParC relatively difficult to understand will be more likely to choose PV over MaxParC.

*Hypothesis 1c. Extreme Complexity or Chance Aversion.*  $S$  will be zero, collapsing across treatment assignment.

Additionally, evidence from behavioral experiments suggests that inequality aversion influences choice in decisions involving multiple actors. Since MaxParC is, by some definitions, the fairer method, it could be more likely to be chosen when there is greater potential for inequality.

*Hypothesis 2a. Inequality Aversion.*  $S$  will be greater for the “high inequality” treatment than the “low inequality” treatment.

On the other hand, the “high inequality” treatment has two different effects on inequality. For each *individual* option, it leads to higher inequality between voters’ payoffs than in the other treatment if that option wins. But the variation of the level of inequality *across* the five options, i.e., the difference in inequality between the most equal and most unequal option, is *lower* in this treatment than in the other. Respondents may thus have fewer inequality-related reasons to prefer MaxParC over PV in the “high inequality” treatment than in the “low inequality” treatment.

*Hypothesis 2b. Inequality Aversion.*  $S$  will be greater for the “low inequality” treatment than the “high inequality” treatment.

*Hypothesis 2c. Inequality Aversion.* More inequality-averse or fairness-loving individuals will be more likely to choose MaxParC than PV.

Finally, as individuals with higher advantage value will more often profit from the majority choice, they should prefer the majoritarian method:

*Hypothesis 2d. Personal Advantage.* Individuals with a higher advantage value will be more likely to choose PV than MaxParC.

We tested our main hypotheses about aggregate choice using a binomial exact test (H1a, H1c) and Fisher’s exact test (H2a, H2b). These analyses take the same primary dependent variable of interest: the share choosing MaxParC. For H2a and H2b, our primary independent variable of interest is the inequality treatment condition.

We analysed individual choices using a logistic regression to predict choice of PV (dependent variable, coded 1/0). The independent variables in these regressions and the hypotheses they address are: (H1b) risk aversion, simplicity preference, and perceived difficulty of MaxParC in three separate regressions; (H2c) egalitarian and pro-social preferences as measured by the Social Dominance Orientation scale; and (H2d) individual advantage value. We included the following standard controls in all regressions: age, education, gender, income and political affiliation. The text for all the questions used in the analyses can be found in the pre-registration. We apply a standard significance threshold of  $p < .05$  for inference in all analyses. We present the statistically significant results in Table 1.

**Results.** Turning to the results of the virtual lab experiment, we find that a significant and sizable fraction of our participants chose to implement MaxParC when given the choice between MaxParC and PV in the free choice rounds (share: 32%, confidence interval  $CI = [29\%, 36\%]$ ,  $p < .001$ ), indicating that about a third of the sample preferred this option despite the presence of chance, additional complexity relative to

PV, and no prior experience with the method. Indeed, over 68% of the sample either strongly agree or agree with the statement “Did you feel that [MaxParC] was fair?” and average earnings were comparable across the two methods: 36.3 points (PV) vs. 35.5 points (MaxParC).

We also find that the fraction choosing MaxParC in the free choice rounds is significantly less than the 44.8% predicted under the assumption of self-interested rational actors. The lower uptake of MaxParC might be due to the perceived difficulty associated with the method. In keeping with our hypotheses, we found that participants who perceive MaxParC to be difficult are significantly more likely to choose PV (standardized regression coefficient  $\beta = .44$ ,  $p = .020$ , see col. 2 of Table 1). Indeed, the fact that we see fewer people choosing MaxParC than expected under self-interested rationality could also be due to people not understanding which voting behavior would be optimal for them. At the same time, they were given 45 s to decide, which in ample piloting seemed to be sufficient for participants to understand the choices.

Contrary to our pre-registered hypothesis, we also found that respondents with higher risk aversion, measured using a self-reported qualitative measure from Falk et al. (2018), were less likely to choose PV ( $\beta = -0.42$ ,  $p = .03$ , see col. 1 of Table 1). All together, these results suggest that the complexity and stochasticity present in MaxParC are insufficient to deter a large fraction of respondents from choosing this method over the deterministic, simpler and familiar Plurality Vote, despite no prior experience with the method. However, they also suggest that some subjects with lower advantage values may not have realized that the welfare-enhancing effects of MaxParC in this experiment (due to its higher likelihood of realizing the socially optimal option) would also tend to be in their own self-interest. Hence additional training and explication may be needed for at least a portion of the population.

We found no evidence that the inequality treatment contributed to the share of participants choosing MaxParC in the free choice rounds. It could be that the effect of the within-option distributive inequality, which was larger by design in the high inequality condition, was harder to detect than the greater inequality across the five options in the low inequality condition. We found no evidence that individual-level inequality aversion, as measured by the items of the Social Dominance Orientation scale (Pratto et al. 1994), predicted choice of voting method. However, we did find that individuals with a higher level of personal advantage in the group—that is, those whose preferences were more likely to be aligned with the majority—were significantly more likely to choose PV over MaxParC ( $\beta = 1.04$ ,  $p < .001$ , see col. 3 of Table 1). In an exploratory analysis of the MaxParC voting rounds, we found that the more risk-averse a respondent was, the more likely they were to give the full consensus option a high rating ( $\beta = 1.65$ ,  $p = .001$ , see *Supplement*, Table S3 for details).

Finally, we find that the average realized payoffs in the 16 rounds of PV and MaxParC that precede the preference elicitation rounds are significantly predictive of vote method choice. Greater earnings in PV relative to MaxParC is associated with a greater likelihood of choosing PV ( $\beta = 1.33$ ,  $p < .001$ , see col. 4 of Table 1), controlling for the advantage value. Similarly, we find that realized earnings in the PV rounds positively predict the likelihood of choosing PV, while earnings in MaxParC are negatively associated with choosing PV ( $\beta = 1.71$ ,  $p < .001$  and  $\beta = -1.71$ ,  $p < .001$ , see col. 5 of Table 1). Finally, we find that these results hold even after including all



**Table 1** Logistic regression models predicting the likelihood of choosing Plurality Voting over MaxParC in the free choice portion of the lab experiment

	<i>Dependent variable</i>					
	(1)	(2)	(3)	(4)	(5)	(6)
Risk aversion	-0.418*					-0.409*
Difficulty		0.438*				0.383
Advantage			1.035***	0.574**	0.740**	0.772**
Pay difference				1.332***		
PV avg pay					1.705***	1.651***
MPC avg pay					-1.707***	-1.715***
Age	0.043	-0.071	-0.003	-0.038	-0.043	-0.053
Education	0.002	0.015	-0.048	0.009	0.004	-0.067
Female	0.807	0.901	1.654	1.254	1.375	1.343
Male	0.740	0.942	1.648	1.145	1.268	1.201
Income	0.196	0.270	0.173	0.187	0.189	0.188
Republican	0.618	0.552	0.451	0.296	0.276	0.272
Democrat	0.615	0.543	0.408	0.199	0.176	0.202
Independent	0.495	0.429	0.307	0.115	0.100	0.133
Observations	616	615	616	616	616	615
R <sup>2</sup>	0.001	0.001	0.002	0.005	0.005	0.005
Adjusted R <sup>2</sup>	-0.014	-0.014	-0.012	-0.012	-0.013	-0.016
Log likelihood	-382.570	-381.016	-369.198	-342.263	-341.877	-336.754

Standardized coefficients from logistic regression models assessing the effect of our primary explanatory variables—risk aversion, perceived difficulty of MaxParC, experimentally assigned advantage value, difference in average realized payoffs (PV – MaxParC)—on the likelihood of choosing Plurality Voting over MaxParC in the free choice portion of the virtual lab experiment. All regressions include the same standard demographic controls: age, measured in integers; education and income, both ordinal 5-pt scales treated here as continuous variables; female, male, republican, democrat, independent and minority, all binary variables indicating whether the category was selected by the respondent (1) or not (0). Note that one respondent did not complete the perceived difficulty question, resulting in one fewer observation. Detailed information and text for all questions used in the analyses can be found in the pre-registration

\* p < 0.05; \*\* p < 0.01; \*\*\* p < 0.001

other significant predictors, suggesting that realized earnings and personal experience with the two voting methods are a strong predictor of choice in subsequent rounds.

## 6 Discussion

Non-deterministic methods can achieve proportional fairness and support consensus, where deterministic methods cannot, without sacrificing total welfare. They also meet other important criteria, such as those necessary for non-manipulability, while employing randomization judiciously. However, the proportional fairness of non-deterministic methods results from their average proportionality over many decisions. Few would recommend electing a government, a consequential decision that happens relatively infrequently, by flipping a coin. In contrast, using coin flips for decisions that happen on a regular basis may be more acceptable because the stakes are often lower and advantages level out over time. While coin flips might achieve proportional fairness, they would not lead to a consensus. The non-deterministic method introduced here (MaxParC) has the benefit of incentivizing compromise in most situations while relegating coin flips to rare occasions where a consensus cannot be found. Indeed, our theoretical and simulation results show MaxParC to be a desirable group decision mechanism, preferable in many ways to Plurality Voting. Voters can reduce or avoid unwanted randomization in MaxParC by rating compromise options positively. Thus, the more risk-averse the voters, the more likely consensus will emerge with MaxParC. Of course, one consequence of radical proportionalism is that stubborn voters who refuse to compromise can still achieve their desired (and potentially extreme) outcomes with a small probability. This means that the options up for vote must not violate any basic rights. This is true of both majoritarian and non-majoritarian group decision methods. History has demonstrated many cases where majorities have voted for options that compromised the basic rights of minorities.

In our behavioral experiments, we found that about one third of voters chose MaxParC over Plurality Voting in the preference elicitation task, suggesting a sizable demand for alternative voting methods—though we did find that perceived difficulty of MaxParC reduced the likelihood that it was chosen. Critically, we also found that a voter's position within their group—specifically, whether they are “advantaged” or hold a majority position—is highly predictive of choosing Plurality Voting. This suggests challenges for the political feasibility of implementing a method that prioritizes fairness but may also be seen to compromise the “will of the majority.” There are of course limitations in the ability of a stylized experiment to capture the complexity of real-world group decisions. In particular, while we operationalized choice options in terms of a monetary allocation, preferences among real-world options are likely to be complex and difficult to represent by a single monetary currency. Additionally, voters are likely influenced by procedural preferences and fairness perceptions.

Finally, while consensus-supporting proportional methods of the type discussed here are well-suited to the governance and management of day-to-day affairs, they could also be applied to more consequential budget allocation decisions or parliamentary elections. The asset distributed by a proportional method need not be the winning probability in a single-outcome decision as in this article, but could be some

other resource to be shared. In that case, the method turns into a deterministic method that outputs a certain allocation of a resource across options. For example, the asset might be parliamentary seats: suppose that instead of one of the common simple proportional methods, the Nash Lottery or MaxParC was used to allocate parliamentary seats to party lists based on voters' ratings of all parties, converting winning probability shares into seat shares using a suitable rounding method. This method might take better advantage of opportunities for consensus without sacrificing proportionality. The seat distribution would, on average, be less fragmented than under existing seat-allocation methods, without sacrificing representation of minorities. Our theoretical, simulation and experimental results thus suggest that communities may be able and willing to avoid methods that result in knife-edge decisions that satisfy only half of voters, as in many recent elections and referenda, by experimenting with consensus-supporting proportional voting methods.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s00355-024-01524-3>.

**Acknowledgements** We'd like to thank three anonymous referees, Marius Amrhein, Markus Brill, Pascal Führlich, Anne-Marie George, Tom Griffiths, Ulrike Kornek, Fabrizio Kuruc, Adrian Lison, E. Keith Smith, Lea Tamberg, Elke U. Weber, and the members of the 'election methods' electronic mailing list (Lanphier et al. 1996) for inspiration and comments. We especially thank William D. Thompson for his help with the Dallinger framework.

**Author Contributions** All authors conceptualized the study, developed theory and methodology, performed the formal analysis, and wrote the manuscript. J.H. developed the software and performed the simulations.

**Funding** Open Access funding enabled and organized by Projekt DEAL. The virtual lab experiment was funded by the grant "Society-Scale Behavioral Simulations through Crowdsourcing, Schmidt DataX Fund, DataX award" (PIs Tom Griffiths, Alin Coman, Simon Levin, Betsy Levy Paluck, Elke Weber, USD 123,928, 9/2019–8/2020.) The authors gratefully acknowledge the European Regional Development Fund (ERDF), the German Federal Ministry of Education and Research and the Land Brandenburg for providing resources on the high-performance computer system at PIK.

**Availability of data and materials** The data that support the findings of this study are available from the corresponding author upon reasonable request.

**Code availability** The simulation programming code that support the findings of this study are available from <https://github.com/pik-gane/vodle/tree/abm>, which also contains open-source code for an online group decision tool based on the MaxParC method presented here.

## Declarations

**Conflict of interest** The authors declare that they have no competing interest.

**Ethics approval** The virtual lab experiment was approved by the Princeton University Institutional Review Board.

**Consent to participate** All participants provided informed consent.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence,

and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Abramowitz B, Elkind E, Grossi D et al (2021) Democratic forking: choosing sides with social choice. In: Algorithmic decision theory: 7th international conference, ADT 2021, Toulouse, France, November 3–5, 2021, proceedings 7. Springer, Berlin, pp 341–356
- Amar AR (1984) Notes choosing representatives by lottery voting. *Yale Law J* 93(7):1283–1308
- Aziz H, Bogomolnaia A, Moulin H (2019) Fair mixing: the case of dichotomous preferences. In: ACM EC 2019—proceedings of the 2019 ACM conference on economics and computation. Association for Computing Machinery, New York, USA, pp 753–781
- Börgers T, Smith D (2014) Robust mechanism design and dominant strategy voting rules. *Theor Econ* 9(2):339–360
- Bouton L (2013) A theory of strategic voting in runoff elections. *Am Econ Rev* 103(4):1248–1288
- Bouton L, Llorente-Saguer A, Malherbe F (2018) Get rid of unanimity rule: the superiority of majority rules with veto power. *J Polit Econ* 126(1):107–149
- Brams SJ, Fishburn PC (1978) Approval voting. *Am Polit Sci Rev* 72(3):831–847
- Brandl F, Brandt F, Seedig HG (2016) Consistent probabilistic social choice. *Econometrica* 84(5):1839–1880
- Carnap R (1947) Probability as a guide in life. *J Philos* 44(6):141
- Carroll R, Lewis JB, Lo J et al (2013) The structure of utility in spatial models of voting. *Am J Polit Sci* 57(4):1008–1028
- Cederman L, Wimmer A, Min B (2010) Why do ethnic groups rebel? New data and analysis. *World Polit* 62(1), 87–119
- Cohen FS (1997) Proportional versus majoritarian ethnic conflict management in democracies. *Comp Polit Stud* 30(5):607–630
- Collier P (2004) Greed and grievance in civil war. *Oxf Econ Pap* 56(4):563–595
- Cross JG (1973) A stochastic learning model of economic behavior. *Q J Econ* 87(2):239–266
- Darmann A (2019) Manipulability in a group activity selection problem. *Soc Choice Welf* 52(3):527–557
- Darmann A, Elkind E, Kurz S et al (2012) Group activity selection problem. In: Internet and network economics: 8th international workshop, WINE 2012, Liverpool, UK, December 10–12, 2012. Proceedings 8. Springer, Berlin, pp 156–169
- Davis JH (1992) Some compelling intuitions about group consensus decisions, theoretical and empirical research, and interpersonal aggregation phenomena: selected examples 1950–1990. *Organ Behav Hum Decis Process* 52(1):3–38
- Dellis A (2010) Policy moderation and endogenous candidacy in Approval Voting elections. In: Laslier JF, Sanver MR (eds) Handbook on approval voting, chap 18. Springer, Berlin, pp 431–454
- Devotta N (2005) From ethnic outbidding to ethnic conflict: the institutional bases for Sri Lanka's separatist war. *Nations Natl* 11(1):141–159
- Dubey P, Shapley LS (1979) Mathematical properties of the Banzhaf power index. *Math Oper Res* 4(2):99–131
- Duddy C (2015) Fair sharing under dichotomous preferences. *Math Soc Sci* 73:1–5
- Falk A, Becker A, Dohmen T et al (2018) Global evidence on economic preferences. *Q J Econ* 133(4):1645–1692
- Fehr E, Schmidt KM (2006) The economics of fairness, reciprocity and altruism—experimental evidence and new theories. In: in Kolm S-C, Mercier Ythier J (eds) Handbook of the economics of giving, altruism and reciprocity: foundations, vol 1. Elsevier, pp 615–691
- Fishburn PC (1974) Simple voting systems and majority rule. *Behav Sci* 19(3):166–176
- Flanigan B, Gözl P, Gupta A et al (2021) Fair algorithms for selecting citizens' assemblies. *Nature* 596(7873):548–552

- Gill J, Gaijdos J (2002) Why does voting get so complicated? A review of theories for analyzing democratic participation. *Stat Sci* 17(4):1–22
- Granovetter M (1978) Threshold models of collective behavior. *Am J Sociol* 83(6):1420–1443
- Harsanyi JC (1973) Games with randomly disturbed payoffs: a new rationale for mixed-strategy equilibrium points. *Int J Game Theory* 2(1):1–23
- Heitzig J, Simmons FW (2012) Some chance for consensus: voting methods for which consensus is an equilibrium. *Soc Choice Welf* 38(1):43–57
- Jones B, Radcliff B, Taber C et al (1995) Condorcet winners and the paradox of voting: probability calculations for weak preference orders. *Am Polit Sci Rev* 89(1):137–144
- Kawai K, Watanabe Y (2013) Inferring strategic voting. *Am Econ Rev* 103(2):624–662
- Kingma DP, Ba J (2014) Adam: a method for stochastic optimization. arXiv preprint. [arXiv:1412.6980](https://arxiv.org/abs/1412.6980)
- Lackner M (2020) Perpetual voting: fairness in long-term decision making. In: Proceedings of the AAAI conference on artificial intelligence. Association for the Advancement of Artificial Intelligence, Washington, DC, USA, pp 2103–2110
- Langhner R et al (1996) Election-methods electronic mailing list. <https://electorama.com/em/>. Last accessed 9 Oct 2020
- Laslier JF (2010) In silico voting experiments. In: Handbook on approval voting, pp 311–335
- Laslier JF, Remzi Sanver M (eds) (2010) Handbook on approval voting. Springer Science & Business Media
- Lewis D (2013) Direct democracy and minority rights: a critical assessment of the tyranny of the majority in the American states. Routledge, London
- May KO (1952) A set of independent necessary and sufficient conditions for simple majority decision. *Econom: J Econom Soc* 20(4):680–684
- McCarty N, Poole KT, Rosenthal H (2016) Polarized America: the dance of ideology and unequal riches. MIT Press, Cambridge
- Merrill S (1988) Making multicandidate elections more democratic. Princeton University Press, Princeton
- Persson M, Esaiasson P, Gilljam M (2013) The effects of direct voting and deliberation on legitimacy beliefs: an experimental study of small group decision-making. *Eur Polit Sci Rev* 5(3):381–399
- Pratto F, Sidanius J, Stallworth LM et al (1994) Social dominance orientation: a personality variable predicting social and political attitudes. *J Personal Soc Psychol* 67(4):741
- Spenkuch JL, Montagnes BP, Magleby DB (2018) Backward induction in the wild? Evidence from sequential voting in the US senate. *Am Econ Rev* 108(7):1971–2013
- Troyan P (2012) Comparing school choice mechanisms by interim and ex-ante welfare. *Games Econ Behav* 75(2):936–947
- Wiedermann M, Smith EK, Heitzig J et al (2020) A network-based microfoundation of Granovetter’s threshold model for social tipping. *Sci Rep* 10(1):11202
- Woodall DR (1997) Monotonicity of single-seat preferential election rules. *Discret Appl Math* 77(1):81–98
- Zakaria F (1997) The rise of illiberal democracy. *Foreign Aff* 76(6):22–43