Egalitarianism vs Utilitarianism in Preferential Voting

Pierluigi Contucci[0000−0002−3577−6189]1 and Alina Sîrbu[0000−0002−3947−7143]2

[1] Department of Mathematics
Alma Mater Studiorum–Università di Bologna, Italy
pierluigi.contucci@unibo.it
[2] Department of Computer Science
Università di Pisa, Italy
alina.sirbu@unipi.it

Abstract. Democratic societies base much of their decisions on voting procedures that involve aggregation of individual votes into a winning solution. While for two candidates majority voting can provide satisfactory results, for three or more candidates the winner depends on the voting method employed. In this chapter we analyse preferential voting, where voting ballots consist of a ranking of candidates. We first study the classical Condorcet criterium introduced to maximise the total satisfaction of voters, i.e. the *utilitarian criterion*. We then complement it with a recently introduced method to minimise the total un-evenness of the rewards, i.e. the *egalitarian dimension*. We show, through targeted examples and analysis of synthetic vote data, that the new dimension may lead to more fair results, and can provide resilience to radical voter opinions.

Keywords: utilitarianism, egalitarianism, voting method, opinion aggregation, preferential voting, Condorcet paradox

1 Condorcet theory of democratic vote

During the French Revolution and especially in the years before, due to the progressive delegitimisation of the King's political power, several intellectuals were advancing proposals to rationalise the steps taken by a group during deliberation. Among them a crucial topic of discussion was the selection of a solution among many alternatives, based on the opinion of the group members. This problem nowadays carries the name of *aggregation* of opinions or votes.

The mathematician Condorcet, in his treaty about the progress of the human spirit [2], in what he calls the tenth epoch, i.e. the future, foresees and hopes for an evolution of the social sciences in the same direction as the hard sciences at his times. He considered exemplary the degree of precision and trustability obtained through the systematic use of mathematics, and he claimed that the same method should be applied to the organisation of society. Among the mathematical areas that are more suitable to achieve such results, probability must surely occupy a prominent place. In the same treaty he explains why, by giving several examples related to the rules to be applied in law and political debates.

In this introductory section we explain the main points of his theory of democratic vote and aggregation of opinions. Such ideas are today still at the foundation of the political sciences. Furthermore, with the enormous development of the internet, they are also used within branch of computer science concerned with sorting objects by relevance, with many applications in indexing and search.

The first paramount observation by Condorcet is the acknowledgment of the high level of complexity of the vote theory from its individual starting point up to the necessary synthesis to create consensus. In particular, he observes that the dichotomic option (yes or no, in favour or against, raised hands) is a funnel too narrow to express an individual opinion. It turns out to be a dramatic limitation of free expression and also easily manipulable in the preliminary stages of the vote. The starting point must therefore include at least a set of choices, options or candidates, that each individual can rank according to their preference. For example, in the case of a set of four candidates A, B, C, D, a vote is a ranking of the candidates, possibly with ties, of the form $A > D = C > B$, or, $D > C > A > B$, or $C = D > A > B$ etc.

This extension of the *space of expression* of the individual vote from dichotomic to multivalued has a precise meaning in mathematics: the local field of Condorcet voting theory takes values on the permutation group, or, if ties are allowed, on the Fubini group. Let us introduce some necessary notation. We will call v_i the vote of i-th voter of a group of N individuals. In general v_i will be a weak ordering of k candidates i.e. an element of the set R_k, the Fubini group. The Fubini numbers are the cardinalities of those sets: $|R_1| = 1$, $|R_2| = 3$, $|R_3| = 13$, $|R_4| = 75$, $|R_5| = 541$ etc. With combinatorial-algebraic techniques one can show that $|R_k|$ grows slightly faster than an exponential, precisely by a multiplicative power-law factor c^k with $c \approx 1.44$. This information about the growth rate is more than a mere technicality. It tells us that if the number of candidates is of the order of the hundreds, like for instance for the problem of ranking the hotels of a middle-sized town, the space R_k is not inspectionable. That means that no computer present or future can span it all because the time needed is well beyond the estimated age of the universe. Problems of this type are called NP-complete [8].

The way Condorcet proposes to *aggregate* the opinions reflects the political ideas of his times. We will exemplify with a concrete example: a high school having to decide where to go on a school trip. If the options are only two, say between Rome and Milan, the decision will turn out to be quite straightforward: by raising hands, the most voted option, the majority vote, is the only one compatible with the democratic principles. But if the options are three or more, hence when the topic has some complexity, new and unexpected effects may appear. Let us say that a class of sixty students must decide if going to London, Paris or Rome. The votes cast are represented in the following table:

30	20	10
Paris	Rome	London
Rome	Paris	Paris
London	London	Rome

which is: 30 students have voted the preference Paris>Rome>London, 20 students voted Rome>Paris>London and 10 London>Paris>Rome. We can then compare the options pairwise, i.e. by computing the number of students who prefer one city over another. We obtain:

– Paris wins over Rome 40 to 20
– Rome wins over London 50 to 10
– Paris wins over London 50 to 10

The end result, the winning ranking, is therefore: Paris>Rome>London.

Let us consider another example:

25	9	12	14
Paris	London	Rome	London
Rome	Paris	London	Rome
London	Rome	Paris	Paris

In this case the pairwise comparison provides:

– Paris wins over Rome 34 to 26
– Rome wins over London 37 to 23
– London wins over Paris 35 to 25

This, clearly, does not admit any winner because a cycle appears in the preferences: Paris>Rome>London>Paris. This is known as the *Condorcet paradox*.

Condorcet proposes a solution to this problem, based on the notion of distance among votes:

$$d(v_1, v_2) \;=\; \text{minimum number of permutations to transform } v_1 \text{ into } v_2.$$

A few examples of distances are:

– $d(A > B > C, B > A > C) = 1$ (swap A with B in the first ranking to obtain the second ranking)

– $d(A > B > C, C > A > B) = 2$ (swap B with C and then A with C)

– $d(A > B > C, C > B > A) = 3$ (swap A with B then A with C then B with C)

Using this measure, we can compute a distance between the result of an election and the vote cast by a voter. If we consider a winning ranking c the i-th voter is distant $d(v_i, c)$ from it. This quantity represents a measure of how unsatisfied with the outcome of the election the voter is. If we apply this to all voters, the total normalised distance from c (the mean distance) is :

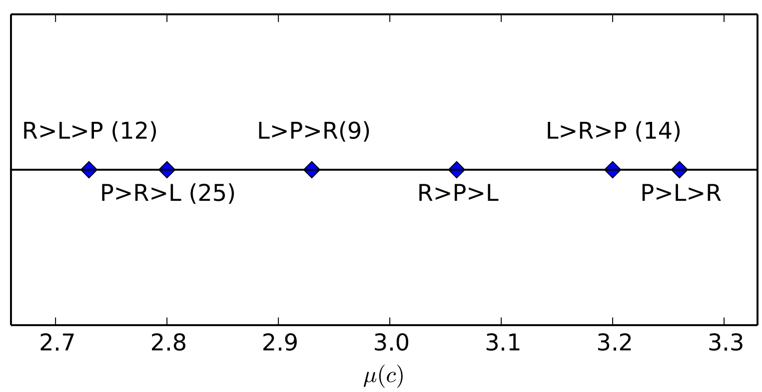
$$\mu(c) = \frac{1}{N} \sum_{i=1}^{N} d(v_i, c) \ .$$

This gives a measure of how unsatisfied all voters are, on average, with the result c.

In order to choose a winner after a vote, we need to choose one suitable c. Condorcet proposes to select as winner the solution c^* that minimises the mean distance from the electorate, μ. Such choice corresponds to choosing the most satisfactory solution and, mathematically, is it obtained with the variational problem:

$$\inf_c \frac{1}{N} \sum_{i=1}^{N} d(v_i, c) \ ,$$

If we consider the second example previously discussed, we can take all possible rankings of the three candidate cities (6 possibilities), and compute for each of them the mean distance from all 60 students. This gives us a mean unsatisfaction level for each possible outcome, and can be represented by the graph:



The Condorcet solution is that with smallest μ, hence in this case it is Rome>London>Paris.

The Condorcet solution, which can be non unique, is the *median* of all the points with respect to the introduced distance and not the *barycentre* among them. This distinction had already been clarified by Toricelli and Cavalieri: the median minimises the sum of the distances while the barycentre minimises the sum of the square distances. In spite of the that, the confusion of the two concepts keeps coming back and sometimes causes harm. In 1919 the United States Census Bureau defined the *population center* of a region using the barycentre instead of the median resulting in a incorrect computation. The mistake was corrected only ten years later by Corrado Gini in [4].

A few final remarks to conclude the section. It was discovered in 2001 that the medieval philosopher Ramon Lull knew already the combinatorial structure of the voting space and also the Condorcet solution [7]. The two contribution are in any case regarded as independent. In Lull's theory probabilistic concepts are completely absent.

The distance introduced by Condorcet is only one possible way to make the Fubini space a metric space. Nowadays we know that those different metrics are classified in equivalence classes and have different impacts on different application fields. It's interesting to note that most of the research in this field are carried inside the tech giant companies like Yahoo, Google and Facebook.

The theory introduced by Condorcet was later refined mathematically [6, 11]. The distance between two votes is also known as the Kemeny distance, while the voting method can be found under the "Kemeny-Young" name as well.

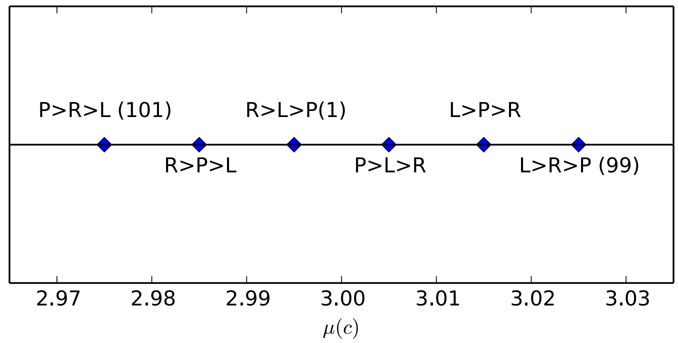
Finally we want to note that although the notions introduced so far are of combinatorial nature (the space of votes), geometric nature (the distance among votes), and analytical nature (the computation of minima), it is indeed the probabilistic nature that is most intrinsically linked to the problem we study: the v_i are, in modern terms, *random variables* describing the macroscopic behaviour of a system composed of N parts (the voters). Condorcet provides a mathematical framework to this problem identifying a solution as a variational problem and opens a new perspective rich of important consequences.

2 A recent development of Condorcet theory

In order to explain a newly introduced idea toward a theory of democratic voting let us consider a different set of votes, this time with a strong polarisation:

101	99	1
Paris	London	Rome
Rome	Rome	London
London	Paris	Paris

Hence, 101 students prefer Paris>Rome>London, 99 students have the totally opposite preference, London>Rome>Paris and one other student prefers London> Rome>Paris. Condorcet theory would only allow to chose from the mean variational principle according to the evaluations of the mean distance (unsatisfaction):



As expected the Condorcet solution turns out to be P>R>L. However, please note the value of μ for the other possible rankings. The median solution is picking up a winning ranking according to the infinitesimal difference of three parts out of a thousand with respect to the second one R>P>L. Is this a good choice?

In order to understand better the question let us go back to the choice among only two alternatives, when we use the majority rule to choose the winner. We know that large majority decisions are appreciated and have a strong stability in time. Instead when the majority rule selects the result by small percentages there is instability and turmoil. Is there a quantity that can measure this type of tension and instability?

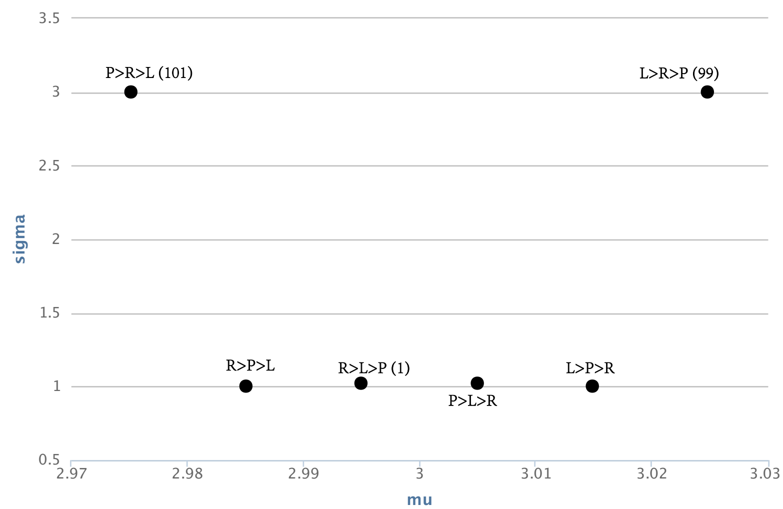
Two examples might clarify the question. Let us consider the vote of 100 individuals choosing among two representatives, A and B. If A receives 95 votes and B 5 the obvious election of A makes 95 people happy and 5 unhappy. Calling $p = 0.95$ the average satisfaction is $2p - 1 = 0.9$, the mean of the binomial distribution. In the case instead in which the two candidates get 51 and 49 votes, the Condorcet solution provides a mean satisfaction of about 0.02. Probability theory provides another important measure, the standard deviation $\sqrt{p(1-p)}$ which quantifies how unevenly is the satisfaction distributed among voters. The lower the standard deviation, the more even is the distribution of satisfaction. The computation of the standard deviation gives, in the first case, about 0.21, while in the second case it is 0.49, a value close to the maximum possible.

It is clear that the standard deviation has a high relevance in many questions of social choice theory because it averages the comparisons among individuals. In economic theory, where social choice is studied, it is well known that personal satisfaction is not only related to personal wealth and its maximization, called utilitarianism, but especially to how ones wealth compares to that of acquaintances, i.e. egalitarianism. The influence of the comparison with respect to the perceived mean has been clarified in the quantitative work [10] of the Economics Nobel laureates Kahneman e Tversky .

We thus proposed [3] to introduce in voting theory a new dimension which is precisely the standard deviation of the distances:

$$\sigma(c) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}[d(v_i, c) - \mu(c)]^2}.$$

This measures the inequality in satisfaction, and allows us to have an extra criterion to select among possibilities, namely an *egalitarian criterion*. We can compute the mean distance μ and the standard deviation of the distances σ for all possible outcomes of the election (all possible rankings) and plot them in 2 dimensions. For the previous example, we obtain:



The figure clearly displays that the selection based only on the *mean* operates on infinitesimal quantities (horizontal axis) and appears to be basically arbitrary with respect to small fluctuations. The vertical axis, however, corresponding to the standard deviation, discriminates much better between possible solutions. Therefore one could consider a different choice, namely the solution R>P>L, which has a standard deviation three times smaller than the one emerging from the Condorcet criterion (P>R>L) and that is likely going to exhibit a higher stability. We have purposely left the concept of stability as a purely intuitive one here. For details please see the original paper [3], which demonstrates through subsampling that points of low standard deviation are more stable with respect to small fluctuations in the votes cast.

3 Egalitarian voting in simulations

To further test the newly introduced method in more realistic settings, we generate synthetic votes from larger populations with various polarisation degrees.

Our method is then applied to the resulting votes. The aim is to understand the role of the egalitarian dimension (σ), how this depends on the polarisation of the population, and how existing *heuristic voting methods* (Schulze, Tideman, Borda, Copeland [1]) compare among themselves with respect to σ.

In order to generate the ranked ballots for each voter, we first generate a set of ratings for each candidate, which we then use to rank them. We fix the number of candidates to $C = 5$ (A,B,C,D,E) and the number of voters to $N = 10000$. A recent analysis of ratings given by voters to real political candidates, in an online experiment [5], showed that, in general, voters tend to rate a few candidates very well, and many candidates very low, with an exponential distribution of ratings between the two extremes. We take this into account and try to reproduce the distribution of ratings observed in this real experiment.

Ratings are distributed in the interval $[-1, 1]$, with a positive rating corresponding to a positive opinion of the candidate. We assume that voters support two opposing parties, we call them Party 1 and Party 2. We consider the candidates A,B,C,D,E, to be ordered by the degree of popularity in the two parties. That means A is the first favoured candidate in Party 1 and E is the favoured by Party 2 voters, while B, C and D are moderate candidates in between the two parties. Each voter gives a rating to each candidate. If a candidate is close to the voter's team, then the rating will be extracted randomly from an increasing exponential distribution that peaks at +1. If, on the contrary, the candidate belongs from the other side of the spectrum, the rating is extracted randomly from a decreasing exponential, peaking at -1. The steepness of the distribution is controlled by a rate parameter which is positive (in the first case) or negative (in the second), and changes from candidates A to E. This results in most ratings with values close to ± 1 and some in between. Figure 1 shows a histogram of all ratings obtained after random sampling, for an example simulation, where 50% of voters are from Party 1 and the rest from the Party 2. We can see that the distribution obtained is similar to that of [5], in that most votes concentrate around the ± 1 values (see Figure 2 in [5]).

The procedure outlined above also allows for simulation of populations with various levels of radicalism. That is, a moderate voter would rate their preferred candidate +1, their least preferred -1, and those in the middle would get intermediate votes. On the contrary, a radical voter would rate +1 some candidates and -1 the rest, with no intermediate ratings for the centrist candidates.

In the following, we generate ratings for candidates when the fraction of voters belonging to Party 1 ranges from 100% to 50% of the population, i.e. from a homogeneous to a polarised population. We consider the situation when voters from the two parties are similar in their radicalism level, i.e. the ratings they give to candidates shift from -1 to +1 in the same way (the rates of the exponentials are the same). From the ratings we generate the ranked ballots, that are then passed through our web application [3, 9] to obtain the 2-dimensional representation of the solution space.

Figure 2 shows the solution space for the case of a completely homogeneous population, i.e. all voters come from Party 1. We can observe that the range

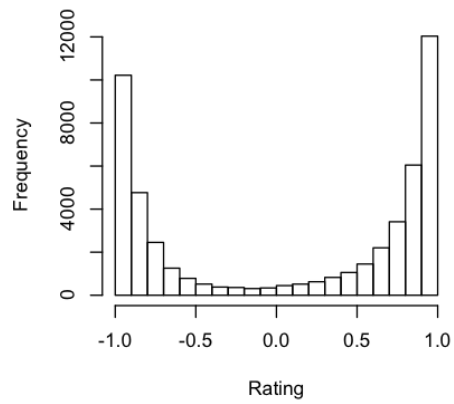


Fig. 1. Histogram of ratings for 5 candidates and 10000 voters (50000 ratings in total).

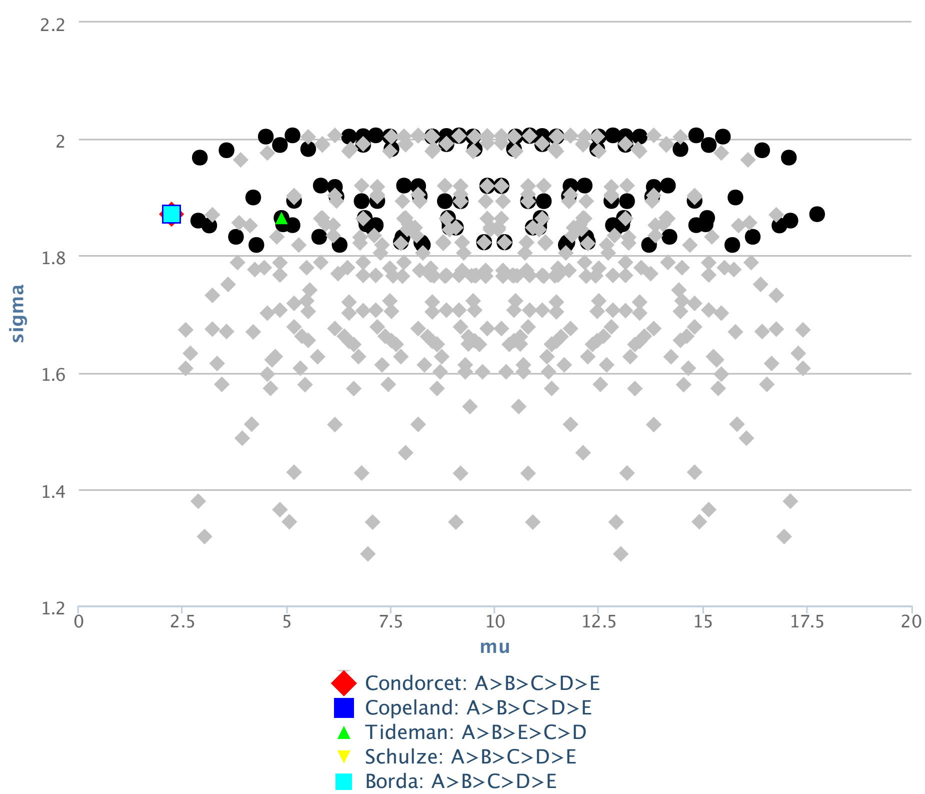


Fig. 2. Utility and egalitarianism for a homogeneous population (100% voters from Team 1). The black dots are solutions without equalities, while the grey diamonds represent the solutions with equalities among candidates.

of the utilitarian dimension (the average) is very wide, while the egalitarian dimension (the standard deviation) has a small range. Hence, in this case, it appears that the utilitarian criterion is enough to distinguish between possible solutions, i.e to select the winner. This because, since all voters are on the same team, their satisfaction with various candidate ratings is similar. Most heuristic voting methods showed in the plot suggest $A > B > C > D > E$ as the winning ranking, which is also the winner by the Condorcet criterion.

We decrease the level of homogeneity of the population, by inserting 25% voters from Party 2, and we show the 2-dimensional space of solutions in Figure 3. We can observe how the egalitarian dimension becomes now much wider, showing that it is most useful when the population of voters is not homogeneous in preferences. However, since a large majority of the population still comes from Party 1, the winner is again $A > B > C > D > E$, as also declared by heuristic methods.

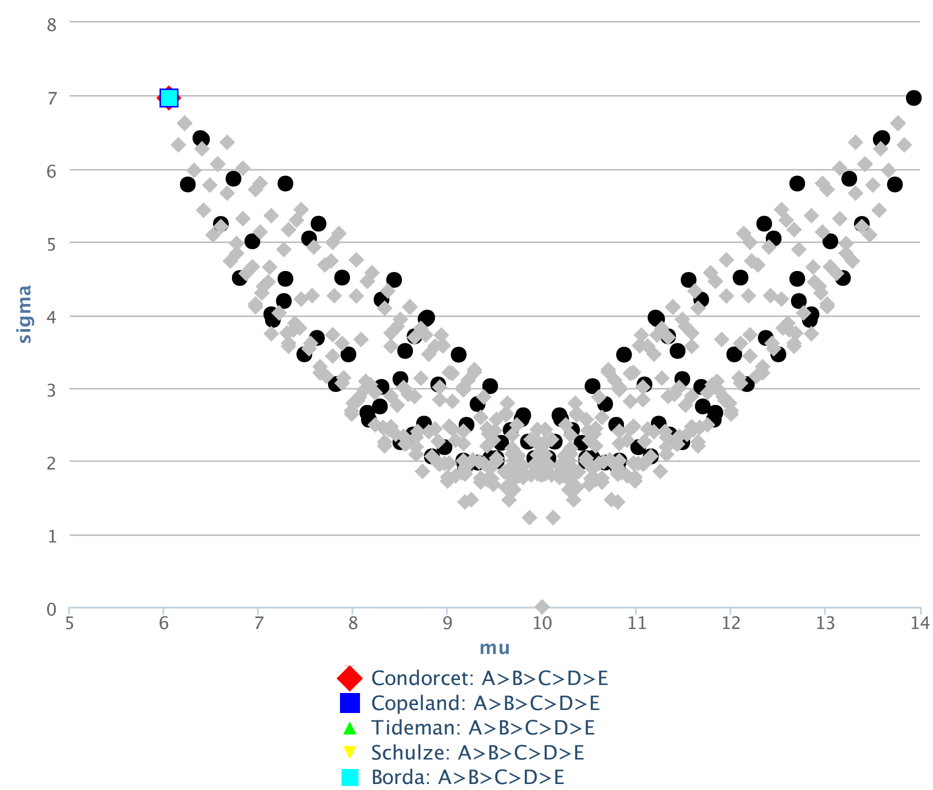


Fig. 3. Utility and egalitarianism for a population with 75% voters from Team 1.

To study the situation of maximum polarisation, we reduce further the fraction of Team 1 voters to 50%, and show the result in Figure 4. We see that here it is the egalitarian dimension that actually dominates the plot. The range of the average distance is very small, which means this criterion has a weak discrimination power, since all possible solutions yield similar average voter satisfaction. Instead our new criterion has a very wide range, hence a very good discrimination power. We thus conclude that the egalitarian perspective is most useful when populations are heavily polarised.

We analyse the figure in detail and see that for all existing heuristic voting methods, the most moderate candidate (C) wins the election, which is very good given that the population is evenly divided between the two teams and voters are similarly radical. The Borda method appears to provide a better candidate ranking from the egalitarianism point of view, while preserving a high average voter satisfaction. We also observe that the point in this area of the plot with lowest σ, i.e. the most egalitarian, is the ranking with equalities $C > A = B = D = E$. This solution has $\sigma = 1.43$ and $\mu = 9.18$, compared to $\sigma = 4.45$ and $\mu = 9.12$ for the Condorcet solution. The most egalitarian solution basically summarises the result saying that, in such a balanced polarised population, candidate C is the best winning choice, while any ranking of the other candidates will decrease egalitarianism.

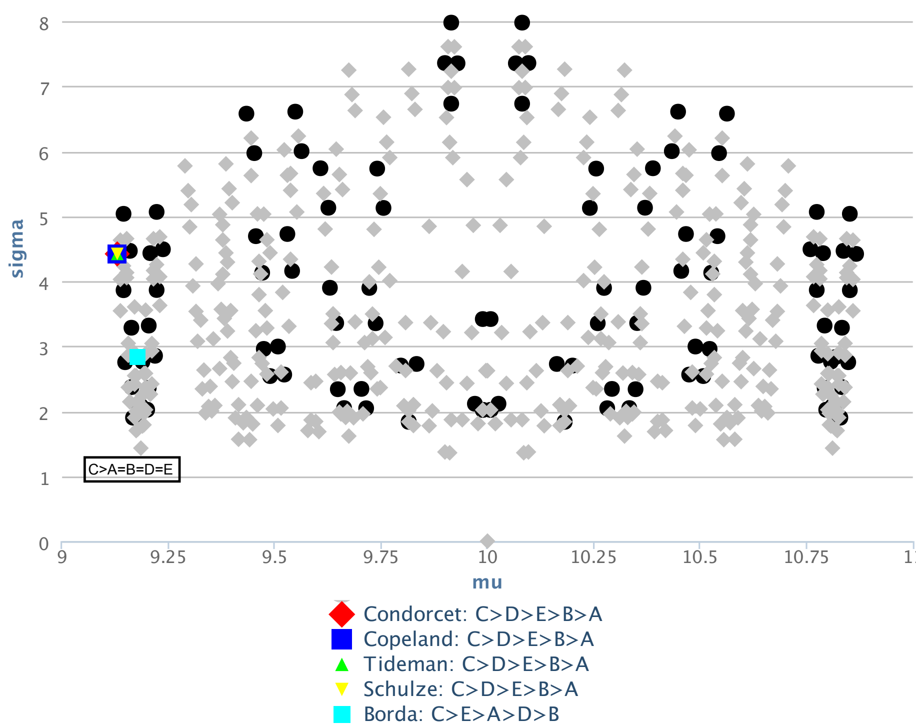


Fig. 4. Utility and egalitarianism for a polarised population (50% voters from Team 1).

We now ask ourselves what happens if the population remains evenly split between teams, but one team (say Team 2) becomes more radical (which in a real setting could correspond to very extremist, outspoken opinions). Figure 5 shows the solution space, with the utilitarian and egalitarian criteria. We can observe that, again, the new criterion has a much higher discriminative power, since the range of values is much wider, while from the point of view of utilitarianism solutions are very close among each other. We also observe that, if we consider the existing heuristic methods, now the winning candidate is E. This means that the more radical team wins, even though the population is evenly split. The Borda solution is again more egalitarian, but the top candidate is still E.

However, if we take into account the egalitarian dimension, we observe that there are solutions with low σ and μ close to the minimum where candidate B wins instead. In fact, if we move from the Condorcet winning ranking, $E > D > B > C > A$, to the most egalitarian ranking without equalities in this area of the plot, $B > E > C > A > D$, we see that μ increases from 9.38 to 9.49 (a factor of 1.01), while σ decreases from 7.58 to 1.75 (4.33 times). We believe this is a much fairer winner, B being more moderate, since the population is evenly split between the two teams. Hence, we conclude that the introduction of this second dimension can make voting more robust to radical opinions.

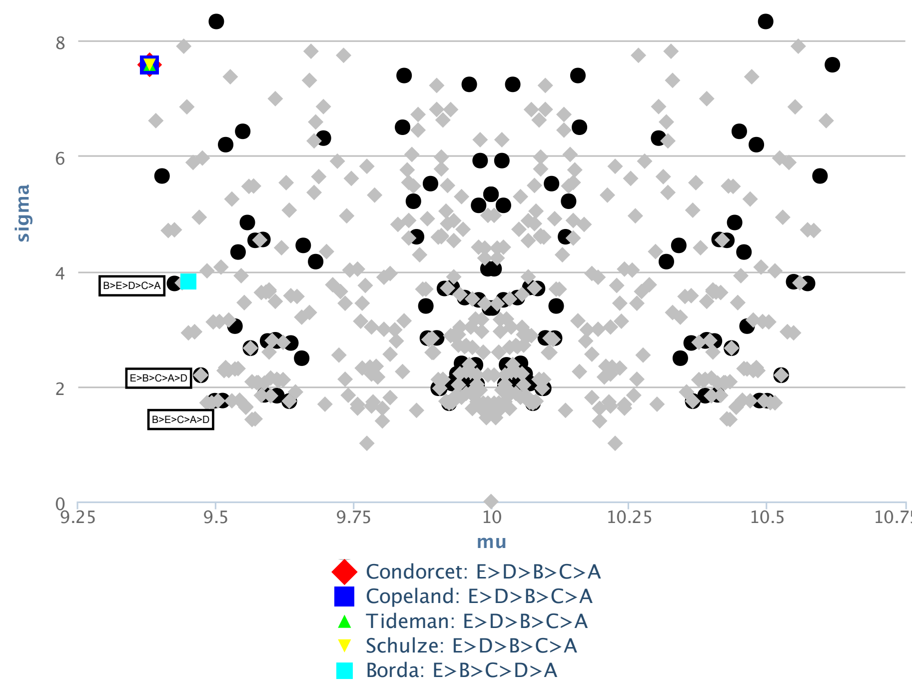


Fig. 5. Utility and egalitarianism for a polarised population (50% voters from Team 1) with radical Team 2 voters.

4 Conclusion

We have reviewed how the concepts of utilitarianism and egalitarianism are both necessary to implement an aggregation criterion to select a winner in a democratic voting process. The former, introduced by Condorcet, guarantees that the total satisfaction of the voters is maximal. The latter ensures that the distribution of the satisfaction is not too uneven.

The two criteria are complementary, and we believe that both are necessary in order to select truly democratic winners. To support this claim, we analysed several scenarios where candidates come from two different parties, and the support of the voter population is distributed in various ways among the two parties. We showed that the second criterion becomes important in case of polarised populations, which is very common both in political but also in other types of social debates. Additionally, we have observed that, when using the utilitarian criterion only, radicalism in the opinions of voters can force the output of the ballots toward their positions. This effect, however, can be removed by the egalitarian criterion.

In cases of polarised populations, it may happen that by optimising utilitarianism the egalitarian dimension is not optimal, and vice versa, i.e. the two criteria are competing. This generates several optimal results, along the so-called Pareto frontier. In these situations it is the policy maker than needs to decide how to weigh the two criteria. Hence, the general landscape that emerges from this investigation is that consensus in social choice theory is not something that can be completely delegated to rules or algorithms, but the policy maker has an important role.

References

1. Borgers, C.: Mathematics of social choice: voting, compensation, and division, vol. 119. SIAM (2010)
2. Condorcet, M.J.A.N.C.: Esquisse d'un tableau historique des progrès de l'esprit humain.. Yves Gravier (1798)
3. Contucci, P., Panizzi, E., Ricci-Tersenghi, F., Sîrbu, A.: Egalitarianism in the rank aggregation problem: a new dimension for democracy. Quality & Quantity 50(3), 1185–1200 (2016)
4. Gini, C.: L'uomo medio. Giornale degli economisti e rivista di statistica 48(1), 1–24 (1914)
5. Gravino, P., Caminiti, S., Sîrbu, A., Tria, F., Servedio, V.D., Loreto, V.: Unveiling political opinion structures with a web-experiment. In: 1st International Conference on Complex Information Systems, COMPLEXIS 2016. pp. 39–47. Science and Technology Publications, Lda (2016)
6. Kemeny, J.G.: Mathematics without numbers. Daedalus 88(4), 577–591 (1959)
7. Llull, R.: Ars notandi, Ars eleccionis, and Alia ars eleccionis (1232 - 1315)
8. Papadimitriou, C.H.: Computational complexity. John Wiley and Sons Ltd. (2003)
9. P.Contucci, E.Panizzi, F.Ricci-Tersenghi, Sîrbu, A.: Rateit web application (2016), http://sapienzaapps.it/rateit.php

10. Tversky, A., Kahneman, D.: Judgment under uncertainty: Heuristics and biases. science 185(4157), 1124–1131 (1974)
11. Young, H.P.: Condorcet's theory of voting. American Political science review 82(4), 1231–1244 (1988)