Democracy 2.1

Karel Janeček, PhD.

August 29, 2016

Is it possible to find a voting system which maintains an equality of voting rights and at the same time gives greater impact to well-informed and engaged voters? Yes, it is possible. Our proposed voting system solves this apparent paradox. In addition to this principal benefit, it also enables voters to rebalance the political spectrum away from extremists and reduce the threat of dishonest or corrupt political actors.

1 Representative democracy

In 1947, Winston Churchill famously declared: "It has been said that democracy is the worst form of government, except all the others that have been tried." Is democracy really as weak, ineffective and unsatisfactory a form of government as it often appears? Might a better system than democracy exist? If not, does it mean that we, as a human society, must endure the ill effects of a system performing far under its potential?

My core conviction is that from a long-term perspective, no other system than democracy will work properly. As soon as the essential principle of equal suffrage is breached, the risk that the social system will eventually degenerate into some form of dictatorship, or another regime which mistreats certain individuals or groups of citizens, is significant. Such a system will necessarily be unstable from a long-term perspective and will lead to major social conflicts.

Selection of high-quality political representation is a key factor for reaching the fairest and most effective functioning of society as a whole. In the case of most voting systems common today, the principle of "one man, one vote" as currently applied has aided the rise of populism and extremism. Relatively easy manipulation and control of the masses by demagogues, growing rates of voter abstention, and the risks of "tyranny of the majority" seem to pose an insoluble problem to 21st-century democracies. These risks are significant: how can we rise to the challenge?

It is additionally important to incentivize civic engagement by allowing citizens who invest their time and effort toward the political and economic development of society (we call them "engaged voters") to see their efforts repaid by a proportionally greater impact on the selection of political representatives.

The basic idea of the voting system presented herein is that the voter gives, in a single ballot, more information than the mere selection of one preferred candidate or party. The overall impact of engaged voters, compared to disengaged or manipulated ones, will thus be higher, without compromising the standard of universal voting rights and equal access of parties and candidates to the political process.

With the power to cast more votes, the electorate as a whole will inherently have greater motivation to consider their choices and deliberate with one another. In short, the system proposed here allows people a better way to express their preferences. By its inherent mathematical properties, this voting system will reinforce civic engagement, electoral participation, and social solidarity.

2 Description of Democracy 2.1

Democracy 2.1 (D21) is neither a fully majoritarian nor proportional system nor a compromise between the two.[1] Its characteristics are as follows:

- 2 seats per district.
- A competing party will nominate one or two candidates.[2]
- Independent candidates may compete subject to other conditions (for example, upon gathering a certain number of signatures from eligible voters).
- Voters may cast up to four plus-votes and up to two minus-votes[3] which they may distribute across all candidates in the district. A voter may give only one vote to any candidate.

For a comprehensive understanding of all effects of the proposed system, let us describe it in the following three steps:

- The effect of more votes (Section 2.1).
- The effect of more seats (Section 2.2).
- The effect of the minus-vote (Section 2.3).

2.1 Effect of more votes

For the simplest explanation, suppose that we have one seat per district and all voters have two plus-votes. The situation in case of two seats and four votes is analogous.

[1] Technically speaking, D21 is a semi-proportional system.

[2] The number of candidates is not constrained; nevertheless, a party would be naturally disincentivized from nominating more than two, as only two seats are available per district.

[3] It might be desired to have only one minus-vote, or under some specific circumstances, no minus-vote at all (for example, where the electorate includes disfavored religious minorities).

2.1.1 Sample situation

To illustrate the strength of D21 we give a model situation. The effects are analogous in practical applications. By "democratic spectrum" candidates, we mean those candidates, regardless of left- or right-wing positions on specific issues, who stand behind the basic guarantees of democratic pluralism, accountability and inclusion. By "extremist", we mean those candidates who stand outside this democratic consensus and seek to use political power to fundamentally alter democratic processes and institutions. By "populist", we mean those candidates who, while nominally part of the democratic spectrum, base their appeal on demagoguery and thus have no natural political allies outside their own party. For the purposes of the illustration, suppose the following candidates are competing:

- One right-wing extremist with 20% of preferences.
- Two right-wing "democratic spectrum" candidates, each with 15% of preferences.
- Two left-wing "democratic spectrum" candidates, each with 15% of preferences.
- One left-wing populist with 20% of preferences.

In the most common majoritarian system, first-past-the-post (FPTP)[4], either the left-wing populist or right-wing extremist would win; under a two-round system, both of them would qualify for the second round of voting. A key characteristic of D21 is an existence of a second vote which each voter may give to another candidate, as follows:

- A supporter of the right-wing extremist is likely to give his second vote to someone from the democratic right or withhold his second vote. Less probably, he will give his second vote to a candidate from the democratic left, and least probably to the populist candidate of the left-wing.
- A supporter of a democratic right-wing candidate will give his second vote most probably to the second democratic candidate of the right wing, less probably to a democratic left-wing candidate, least probably to the right-wing extremist.
- Supporters of left-wing candidates will behave accordingly, with designations of "right" and "left" reversed.

The results of voting can be expected as follows:

- The right-wing extremist will get just over 20% of votes.

[4] In a first-past-the-post or winner-takes-all election, the winning candidate is the one who has received more votes than any others.

- The right-wing democratic candidate will get just over 30% - he will get votes from his supporters (15%) and most votes from supporters of the second democratic right-wing candidate (15%), and some second votes from supporters of the right-wing extremist.
- Similarly, the left-wing democratic candidate will get just over 30%.
- The left-wing populist will get just over 20%.

The analysis shows that in our model one of the candidates of the "democratic spectrum" will be elected. In standard voting systems non-extremist candidates must compete with one another in a zero-sum contest for votes, meaning that right-wing or left-wing extremists or populists are more likely to win or advance to the second round to the detriment of candidates with broader appeal. A key characteristic of D21 is the fact that candidates with broader appeal are less likely to attack one another to their mutual detriment.

2.2 Effect of more seats

An additional weakness of majoritarian, single-winner districts is their suboptimal representativity – small parties can win a very significant number of votes without carrying a single legislative seat. The solution lies in the election of more than one candidate per voting district. In our proposal, each district will elect two winners. The additional seat, combined with the effect of more votes, has a significant and positive impact on representativity. Note that a smaller party has the option of nominating just one candidate, with the goal of attracting additional (third or fourth) votes from the supporters of larger or more well-established parties. A strong candidate nominated by a small party, or even an independent candidate, will thus have a far better chance of winning than under FPTP rules. As a result, new parties and ideas will have far greater opportunities to compete and succeed in the democratic process. Additionally, because of the effects of more seats and more votes, negative campaigning by any one candidate against any other will be less efficient, and thus much less widespread.

We expect that in a parliament or legislative body elected under the D21 system, there will be more smaller parties represented than under a typical proportional system. At the same time, a strong party with broad electoral appeal might become even stronger, thanks to the "majoritarian" effect of our system, than it would under a proportional system. What kind of party will weaken? Medium-size parties, especially those with more narrow electoral appeal, are likely to win fewer seats under D21 than under a typical proportional system.

With D21 we thus expect both higher representativity and innovation than under proportional voting systems, and, because of a more collaborative and less negative electoral environment, an easier process of government formation. The electoral system will at once be more competitive and dynamic, and the formation of governments easier and more stable, than is the case under proportional voting systems.

2.3 Effect of the minus-vote

A final innovation is the third feature, the "minus-vote," which allows voters an even wider scope to express their preferences by letting them designate a candidate they do not want to see elected. A minus-vote and a plus-vote have the same absolute weight (-1 or +1). In itself, the availability of a minus-vote will probably result in a significantly higher rate of voter participation, especially in a climate of public skepticism or disapproval of current political leaders. Moreover, the existence of minus-votes further diminishes the electoral strength of populists and extremists. Minus-votes provide an important means to filter notoriously corrupt and criminal actors from a political system where they are too often protected, for example by "hiding" in a party list and benefiting from the goodwill of voters toward their party and its other candidates.

To illustrate this fact we assume a simple model with one voting district with two parties nominating their candidates. We denote the "right-wing" candidate as R and the "left-wing" candidate as L. Each party will nominate two candidates, one corrupt (-) and one honest (+). We will mark them as L^+, L^-, R^+, R^-. In an ideal case, the honesty of any individual candidate would be a decisive criterion for voters, and minus-votes would be given to corrupt candidates, with winners L^+ and R^+. In our model, we assume the worst possible scenario, where all voters of each party prefer the corrupt candidate of their own party to the honest candidate of the opposing party and, at the same time, they are capable of so-called tactical voting. Tactical voting means that voters will distribute their minus-votes in a manner intended to benefit their preferred candidate rather than register disapproval of another; thus, tactical voters may be willing to cast minus-votes against the honest candidate of the opposing party rather than the corrupt candidate of the opposing party, if they judge this to benefit their own preferred candidate.

Even in this worst-case scenario, the presence of minus-votes acts to purify the system from corrupt actors. The honest candidates L^+ and R^+ are elected as soon as the difference between the size of the right-wing and left-wing electorate is not too great. Specifically, L^+ and R^+ win as soon as the ratio between the number of right-wing voters n_r and the left-wing voters n_l is more than 3 to 4 and less than 4 to 3. In the case that voters do not vote tactically, i.e. where they give their minus-votes to the corrupt candidate of the opposing party, the purifying effect happens for a ratio of votes between 1 to 2 and 2 to 1. See the Attachment for a proof.

The conclusion above holds that both of the honest candidates win as soon as $n_r/(n_l + n_r) \in \left(\frac{3}{7}, \frac{4}{7}\right)$. Without the existence of a minus-vote, both candidates of one of the two parties will win. Therefore, the minus-vote brings significant purification of political representation. A real situation is of course more complicated than the model. Note, however, that the model assumption are very "pessimistic" in terms of preferences and tactics of voters.

3 D21 vs. other voting systems

The well-known Arrow's Impossibility Theorem states that, with at least three distinct alternatives, no voting system can convert the preferences of individuals into a community-wide ranking which fulfills the following natural criteria:

- *unrestricted domain* – all preferences of all voters are allowed,
- *non-dictatorship* – no single person decides by himself or herself,
- *Pareto efficiency* – it is not possible to make one individual better off without making anybody else worse off, i.e., if everybody prefers A over B, then the voting systems must also rank A over B,
- *independence of irrelevant alternatives* (IIA) – the ranking of A over B should depend only on individual preferences between A and B.[5]

3.1 D21 and the Condorcet winner criterion

We can examine different social choice criteria individually and discuss the conditions under which each may be violated. Under this approach, D21 performs extremely well. Take for example the very restrictive *Condorcet winner* (CW) criterion that is usually considered desirable. CW requires that if there exists an alternative that would win a head-to-head contest between itself and any other choice, the voting method must always choose that alternative.

The CW criterion is not satisfied by most voting systems. For the most common voting systems such as FPTP or two-round majoritarian system, a violation appears similar to the example provided in Section 2.1, revealing major weaknesses of FPTP and similar majoritarian systems. (A commonly cited example is the French presidential election of 2003, when the likely Condorcet winner, Socialist Lionel Jospin, was eliminated in the first round of voting.)

D21 does not satisfy CW for a fundamentally different reason than FPTP. Consider three alternatives A, B, C, where $A \succ B \succ C$ for $n+1$ voters, and $B \succ C \succ A$ for n voters, $n \in \mathcal{N}$. The Condorcet winner is A, while the D21 winner is B. B is also the winner in the *Borda count*[6] voting system for $n \geq 2$.

Here we see that the Condorcet winner does not maximize overall social utility. When candidate A wins, all $n+1$ voters are gratified to have their first choice selected, whereas for all n voters, their least-favorite candidate has won! Under D21 (and Borda count), which selects candidate B, n voters now have their first-choice candidate, $n+1$ voters have their second choice, and no voter

[5] We will see further that, unlike the previous three, IIA might not be a desired criteria from the same reason as the Condorcet winner criterion.

[6] In the Borda count system, devised in 1770 by a French mathematician of that name, voters rank candidates in order of preference. Each candidate receives a number of points corresponding to the number of candidates ranked lower that he or she. A candidate with the most points is the winner. Though Borda count proves to be a more consensus-based voting system than FPTP, it does not benefit from the effect of more votes.

has their least-favorite choice. Thus, from a social utility (or merely common-sense) point of view, the choice of B is more rational, especially for larger n. This simple example illustrates that the CW criterion does not always lead to the most reasonable result, and D21 reflects this insight.

Note that this example also violates the IIA from Arrow's theorem since candidate C is the *irrelevant alternative* that changes the voting result for D21 (and Borda count). And yet, the change is desirable from the perspective of overall social utility.

A similar analysis can be done for wide range of other social choice criteria.

3.2 D21 vs. Instant-runoff voting

Arguably the most sensible voting system applied in political voting in practice is the *Single transferable vote (STV)* system. When STV is used for single-winner elections it is known as instant-runoff voting (IRV) (or occasionally "alternative vote", "transferable vote", or "preferential voting"). We show by an example the weakness of IRV compared to D21.

Suppose that we have four candidates that we denote as I (intellectual), C (consensual), P (populist), and E (extremist). By "intellectual", we mean to suggest a candidate who may have exceptional personal qualities, but for reasons of background or style appeals to a smaller, more highly educated set of voters; by "consensual", we mean a candidate who may support the same policies as the "intellectual" candidate but who appeals to a broader electorate (again, for reasons of style or background). "Populist" and "extremist" follow the definitions given in Section 2.1.1 above.

We have four types of voters n_I, n_C, n_P, n_E with the following ranks of preferences:

Type of a voter	Preferences
n_I	$I \succ C \succ P \succ E$
n_C	$C \succ P \succ I \succ E$
n_P	$P \succ C \succ I \succ E$
n_E	$E \succ P \succ C \succ I$

We assume

$$\begin{aligned} n_I > & n_C > n_P > n_E, \\ & n_C < n_P + n_E, \\ n_I < & n_C + n_P + n_E. \end{aligned}$$

Under D21, with two votes per voter, candidate C is a clear winner, getting either the first or second vote of every voter except those with the most extreme preferences. What about IRV? Since no candidate receives more than half of all votes, the weakest candidate E drops and the second preferences of extreme voters n_E are counted as votes for P. This illustrates a key flaw of IRV, where supporters of the weakest candidate may strongly influence the result – more than those whose first and second choices are for more consensual candidates.

This core weakness of IRV is apparent in the scenario given above. In the second calculation, the consensual candidate C is eliminated since $n_P+n_E > n_C$. The second choice of extremist voters together with the first choice of populist voters decide, while the preferences of intellectual voters have no effect since C is the second choice for them and the I candidate is still in the running. With C eliminated and $n_I < n_C + n_P + n_E$, the contest has become a "battle" where demagogy defeats intellect, a result suggestive of an anti-intellectual or a pro-populism mood in society. Note that intellectual voters had no effect on the result at any point in the calculation even though they are the largest group! In fact, the result would be different with C winning if we decrease the number of intellectual voters so that $n_C > n_I > n_P + n_E - n_C$, which again shows the possible logical inconsistency of IRV.

For a specific illustration, imagine a scenario in which $n_I = 8$, $n_C = 4$, $n_P = 3$, $n_E = 2$. The results for D21 voting would be (with a slight abuse of notation) $C = 15$, $P = 9$, $I = 8$, $E = 2$ with candidate C as the clear winner. The IRV algorithm first eliminates E, then since $4 = n_C < n_P + n_E = 5$, the clear D21-winner C is eliminated, with the result that P wins as $8 = n_I < n_C + n_P + n_E = 9$.

Our analysis illustrates why D21 is qualitatively a superior voting system to STV or IRV, despite the substantial merits of those systems.

3.3 D21 vs Approval voting

Approval voting (AV) has a strong theoretical appeal. Like D21, it benefits from the consensus-producing effect of more votes per voter. AV probably selects the candidate with highest overall utility more often than other systems except D21. Here the conceptual difference between "highest consensus" and "highest overall utility" is that whereas "consensus" simply measures the number of voters who are satisfied or dissatisfied with a given outcome (a binary "yes/no" measurement), "utility" factors in the strength of satisfaction or dissatisfaction with that outcome, based on the preferences of each voter. The "highest consensus" candidate may often, but not always, be the candidate producing the highest overall social utility as well.

The central weaknesses of AV is its tendency merely to favor candidates who offend the fewest number of voters. These candidates may thus maximize consensus (in terms of the number of voters who favor rather than disfavor them), while not at all maximizing overall utility (because their support from these voters is relatively tepid). Thus AV tends to result in the selection of a "merely inoffensive" candidate to win, one whom many voters are "fine with" despite not taking any strong positions or showing noteworthy qualities of leadership.

By optimizing the number of plus-votes, D21 reduces the likelihood of this outcome. Furthermore, given a non-trivial number of candidates, the "scarcity" of votes in D21 as compared to AV makes *tactical voting* less likely. Under D21, the *Burr dilemma*[7] is also less likely. For example, the relative "scarcity" of votes

[7] When two candidates, A and B, appear to be the "front-runners", voters are strategically

under D21 makes people less likely to consider another "merely inoffensive" candidate to be a competition for their favorite candidate, and thus are less likely to resort to tactical voting.

We further expect that D21 is likely to increase voter participation significantly more than AV. In the case of AV a voter has no incentive to prioritize any candidate over any other – there is no limit to the number of candidates he or she can select, and no way to distinguish strong from weak support. A limited number of votes thus strengthens the "gaming" aspect of voting, motivating voters to carefully consider and discuss their limited set of choices, including the minus vote, to determine how best to distribute their votes.

4 General definition of D21

Suppose we want to choose W winners out of $T \geq 4$ candidates. A voting system is Democracy 2.1 (D21) if and only if:

- Each voter is allowed to cast up to $P \geq W$ plus votes and up to M minus votes, where $P \geq 2M$ (i.e., number of plus votes has to be at least twice as large as the number of minus votes), and $P \leq T/2$. In most cases, we recommend $P \geq 2W$ (optimizing the *effect of more votes*) and $P \leq T/3$.
- Each voter can cast no more than one vote for any candidate.
- The number of plus-votes cast by each voter must be at least twice the number of minus-votes cast.
- Each vote has the same absolute weight (+1 or -1). The W candidates receiving the greatest sum of all votes win.

5 Conclusion

On a practical basis, it would be difficult to conceive of a scenario under which D21 (with or without the minus-vote feature) would yield a result inconsistent with the maximum overall social utility. D21 offers a new level of theoretical consistency and correctness, and at the same time has already proven itself to be as promising in practice as it is conceptually. D21 is a voting system for the 21st century, as revolutionary, we argue, as an upgrade from DOS to Windows. Ultimately, Winston Churchill was mistaken – because he only knew of Democracy 1.0. All great ideas must evolve to survive, and we believe that D21 represents a next step toward the fulfillment of democracy's promise for humankind.

---

motivated to approve one but not the other. This generates a "vote split" between A and B, which could permit a far-less-favored candidate C to win – as was the case in the U.S. election of 1800, when the less-favored Aaron Burr benefited temporarily from a deadlock between the more-favored John Adams and Thomas Jefferson. Of course, this theoretical weakness is still small compared to the undesirable properties of first-pass-the-post or similar systems.

Effect of minus vote (proof)

Let's denote the utility of the left-wing voter gained by electing the candidate L^+ as $U(L^+)$. The utility of the left-wing voter gained from all individual candidates is then the following:

$$U(L^+) > U(L^-) > U(R^+) > U(R^-).$$

For the right-wing voter, it is similar:

$$U(R^+) > U(R^-) > U(L^+) > U(L^-).$$

Let's also mark the number of right-wing voters as n_p and the number of left-wing voters as n_l and the number of all voters as $n = n_l + n_p$.

Without the loss of generality, let us presume that there are more voters of the left-wing than voters of the right-wing, i.e. $n_l > n_p$. (In the inverse case, all consequent considerations are valid with the conversion of indexes l and p.)

We will analyze two possible scenarios:

1. The voter gives two plus-votes.
2. The voter gives two plus-votes and one minus-vote.

In both cases the two candidates with the highest summation will advance (plus-vote counting as 1 point, minus-vote as -1 point).

Scenario without minus-vote

In the case that the voter does not distribute his minus-votes, the result is evident because all voters vote unanimously. We can see results in the Table 1. Because $n_l > n_p$ both left-wing candidates will advance.

Candidate	Votes
R^+	n_p
R^-	n_p
L^+	n_l
L^-	n_l

Table 1: Number of votes in the model without minus-vote.

In the inverse scenario, where right-wing voters outnumber their left-wing counterparts, both of the right-wing candidates will advance.

Let us remind that n is the number of all voters. Let's mark the share of right-wing voters as q and we can express it as

$$q = \frac{n_p}{n}.$$

The share of left-wing voters is complementary up to 1 (that is, $1 - q$) and for this case it holds

$$1 - q = \frac{n_l}{n}.$$

Then we can express the result in this model depending on the share of voters of the right-wing q in the Table 2.

Right-wing share	Advanced
$q < \frac{1}{2}$	L^+ a L^-
$q > \frac{1}{2}$	R^+ a R^-

Table 2: Results of the model without minus-vote.

Model with minus-vote

In the case that the voter distributes one minus-vote, the analysis of the situation is more complicated. Let us remind that the voter prefers criterion of unanimous voting to the criterion of corruption. In this case, the voter divides his two plus-votes between two candidates of his party. It is obvious that he will give his minus-vote to the candidate of the second party – however, it is not necessarily clear to which candidate. In the case that the voter doesn't vote tactically, he will give his vote to the corrupted candidate. We can see the resulting number of votes in the Table 3.

Candidate	Votes
R^+	n_p
R^-	$n_p - n_l$
L^+	n_l
L^-	$n_l - n_p$

Table 3: Number of votes in the model with minus-vote without tactical voting.

The first one elected is the candidate L^+. In the case of the second candidate it depends whether $n_l - n_p < n_p$. This inequation is an equivalent to $n_p/n_l > 1/2$. If we relate inequations to the share of the right-wing q, we will get

$$q > \frac{1}{3}.$$

If such inequality is fulfilled, the candidate R^+ advances as the second one. If it is the other way round (left-wing strongly prevails), the second elected representative will become the candidate L^-.

The results of the vote is described in the Table 4.

Right-wing share	Advanced
$q < \frac{1}{3}$	L^+ and L^-
$\frac{1}{3} < q < \frac{2}{3}$	L^+ and R^+
$q > \frac{2}{3}$	R^+ and R^-

Table 4: Results of the model with minus-vote without tactical voting.

In the worst case, each voter will use tactical voting (which is, however, not very realistic). In this case, the best option for supporters of the left-wing party is to choose the so-called "dominant strategy",[8] i.e. to distribute minus-votes equally between both two candidates of the right-wing. To do so, the left-wing voters weaken both opposing-party candidates equally, to the benefit of their corrupted candidate. The dominant strategy for the right-wing is to give all minus-votes to corrupted candidate of the left-wing and therefore maximize the chance of their own honest candidate. We can see resulted number of votes 5.

Candidate	Votes
R^+	$n_p - n_l/2$
R^-	$n_p - n_l/2$
L^+	n_l
L^-	$n_l - n_p$

Table 5: Number of votes in the model with minus-vote in the worst case.

The first elected candidate is obviously L^+. Right-wing candidates have approximately the same result of votes. One can assume that the candidate R^+ will get somewhat fewer minus-votes and will have better result than candidate R^-. The right-wing candidate R^+ is elected when $n_p - n_l/2 > n_l - n_p$, which is an equivalent to $n_p/n_l > 3/4$ and expressed in relation to the share of right-wing voters

$$q > \frac{3}{7}.$$

General results are described in the Table 6.

Right-wing share	Advanced
$q < \frac{3}{7}$	L^+ a L^-
$\frac{3}{7} < q < \frac{4}{7}$	L^+ a R^+
$q > \frac{4}{7}$	R^+ a R^-

Table 6: Results of the model with minus-vote in the worst case.

[8] In game theory, a dominant strategy is the strategy which brings the player the best possible result.