



# STAR Voting, equality of voice, and voter satisfaction: considerations for voting method reform

Sara Wolk<sup>1</sup> · Jameson Quinn<sup>2</sup> · Marcus Ogren<sup>3</sup>

Accepted: 24 December 2022

© Springer Science+Business Media, LLC, part of Springer Nature 2023

## Abstract

The vote itself forms the foundation of our political discourse, and how we vote may be the single most important decision in a representative democracy. In the pursuit of more fair and representative elections, we explore common electoral pathologies like vote-splitting, Favorite Betrayal, and the center-squeeze Spoiler Effect, and discuss criteria and methodology for evaluating voting method quality. We introduce STAR Voting for consideration in place of Plurality and Instant Runoff Voting for governmental elections, and compare it to other leading voting method proposals. To evaluate voting method accuracy and strategy resilience, we present the metrics Voter Satisfaction Efficiency (VSE) and Pivotal Voter Strategic Incentive (PVSI). We find that compared to Plurality Voting, all methods tested substantially reduce the need for voters to consider candidate electability but that STAR Voting, Smith/Minimax, the Condorcet method included, and Approval Top Two go further to ensure an equally weighted vote. We conclude that STAR Voting in particular offers a viable, competitive, and compelling proposal for more representative elections.

**Keywords** Voting system · Expected social utility · One person one vote · Strategic voting · Cardinal preferences · Ranked choice voting

**JEL Classification** D72 · D63 · D71 · C70

---

Sara Wolk and Jameson Quinn have contributed equally to this work.

---

✉ Sara Wolk  
sara@equal.vote

Jameson Quinn  
jameson.quinn@gmail.com

Marcus Ogren  
marcus.ogre@gmail.com

<sup>1</sup> Eugene, OR, USA

<sup>2</sup> Cambridge, MA, USA

<sup>3</sup> Ramat Gan, Israel

## 1 Introduction

Voting methods exist at the very core of our political discourse and government, affecting electoral outcomes and incentives for voters, candidates, and parties, and thus shaping our society. In this paper, we present a modern voting method, STAR Voting, discuss its properties, and give some results from statistical models comparing its utilitarian outcomes and strategic incentives to other prominent voting methods.

All voting methods consist of two parts: a ballot, and a process for the tabulation of those ballots. STAR Voting is intended to optimize both of these components. In this paper, we lay out the rationale for the five star ballot as a voter interface, and the Score Then Automatic Runoff (STAR) algorithm for measuring the strength and breadth of support for each candidate in order to find majority preferred winners who represent the will of the people.

## 2 Presenting STAR Voting

STAR Voting is an alternative voting method in which voters score candidates from zero up to five stars, as shown in Fig. 1. The name STAR is both a reference to the five star ballot itself, and an acronym for Score Then Automatic Runoff, which describes the two-round tallying process:

- *Scoring Round*: All scores for each candidate are totaled and the two highest scoring candidates advance.
- *Automatic Runoff*: In the runoff, each ballot is counted as one vote for the finalist who was scored higher on that ballot. The finalist who was preferred by more voters wins.

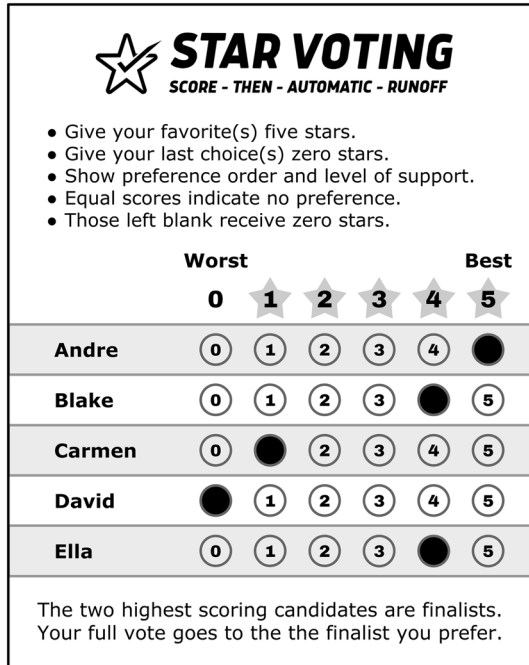
STAR Voting was invented in 2014 with the objective of better delivering on the underlying goals of voting reform advocates, while addressing serious issues with Plurality Voting and limitations with leading reform proposals like Top Two Runoff, Score Voting, and Instant Runoff Voting (IRV).

The five star ballot allows voters to express not only preference order, but also equal preference and preference strength. As a result, the star ballot not only conveys more information than a choose-one ballot, it also conveys more than a ranked ballot. On a strict ranked ballot (where equal preferences are not allowed) a voter's second choice could be as good as the voter's favorite or almost as bad as their last choice. In contrast, even in STAR elections with a large field,<sup>1</sup> there's no need to limit the number of candidates, limit the available levels of voter expression, prohibit equal rankings, or present voters with an unwieldy ballot. Eliminating these

---

<sup>1</sup> For example, the 2020 elections for Democratic presidential delegates from Oregon (The Equal Vote coalition, 2020).

Fig. 1 A STAR Voting ballot



rules reduces the probability of invalid (“spoiled”) ballots, which studies suggest may be increased under Instant Runoff Voting (Neely & McDaniel, 2015).<sup>2</sup>

Before the 0–5 scale was selected for STAR Voting, early simulations on Voter Satisfaction Efficiency (Quinn, 2017) were used to confirm the hypothesis that this level of granularity does in fact lead to significant gains in terms of more representative outcomes compared to less expressive scales. The simulations also confirmed that increasing the ballot scale further yielded diminishing returns. Even with large fields of candidates, the 0–5 star scale offers a high degree of resolution and granularity for voter choice, without exceeding the upward bounds of cognitive load.<sup>3</sup>

The five star rating, well-known from online surveys, is very similar to the five star ballot and offers a familiar interface that has become a leading option for collecting detailed public opinion data. When tabulating five star ballots, both scores and runoff votes are totalled using addition, which allows STAR Voting to be run on existing voting machines in most cases. Tabulation by addition also means that STAR Voting ballots are summable and don’t require centralized tabulation, unlike Instant Runoff Voting (Maine Supreme Judicial Court, 2017). STAR results can be totaled at the local level with accurate reporting of preliminary results as ballots

<sup>2</sup> While some ranked ballot methods like Smith/Minimax allow equal rankings, in practice, most do not.

<sup>3</sup> There have been many psychometric studies of the optimal number of options for scaled responses; see for example Simms et al. (2019). General consensus is that measures such as reliability are relatively stable up to about 6 or 7 options, with cognitive load becoming a limiting factor after that. For this reason STAR Voting employs the 0–5 star rating system.

come in. This makes the five star ballot itself a compelling option for increasing voter voice in elections.

The second half of the STAR Voting method is the ‘automatic’ Top Two runoff. In the runoff, whether or not your favorite can win, your one full vote goes to the finalist you prefer. Ballots that scored both finalists equally count as a vote of ‘no preference’ in the runoff. The STAR Voting runoff identifies majority-preferred winners whenever possible, while voters who oppose both finalists can still, at the least, help to prevent their worst-case scenario. In STAR Voting, the candidate with the most votes wins (as in Plurality), as required by election laws in states and jurisdictions.<sup>4</sup>

While this paper focuses on single-winner voting methods, good multi-winner voting methods are also important for comprehensive electoral reform. There are two ways of using STAR Voting and the five star ballot for multi-winner elections. One option is to repeat the single-winner Score Then Automatic Runoff process until all positions have been filled. The other option is to use Proportional STAR Voting (STAR-PR), which employs the same five star ballot as single-winner STAR but tallies those ballots using a proportional algorithm to ensure that factions or parties that have a quota worth of support can win a proportional number of seats. Evaluating and choosing between these options requires tools that are beyond the scope of this paper.

### **3 Pass/fail criteria, vote-splitting, the ability to vote your conscience, and inequality in the vote**

Comparisons of voting methods have historically largely focused on evaluating methods according to various pass/fail criteria, even though many desirable electoral goals are mutually exclusive or inversely correlated (Arrow, 1950). Despite the obvious limitations of this lens, many advocates frame the conversation around the criteria which their preferred voting methods pass—to the exclusion of other criteria and common sense considerations like representative accuracy. Predictably, this black and white approach has proven to be more divisive than constructive.

Two criteria which are both highly regarded but are inversely correlated are Favorite Betrayal and Later No Harm. The Favorite Betrayal (FB) criterion is about preventing the need for the all-too-familiar strategy where rather than throwing away their vote on a candidate they know can’t win, a voter will vote for the candidate on their side who seems the most electable. Voting methods which pass Favorite Betrayal solve this problem, requiring that a method will never incentivize giving one’s favorite any less than maximum support.<sup>5</sup>

<sup>4</sup> Instant Runoff has been found to be unconstitutional in Maine due to noncompliance with win by plurality and lack of summability (Maine Supreme Judicial Court, 2017).

<sup>5</sup> Favorite Betrayal is defined as: “holding other votes constant, a voter should never have to give their favorite less than the maximum rating in order get a higher-utility outcome.” See Appendix A.

The Later No Harm (LNH) criterion is effectively the opposite of Favorite Betrayal. Saying nothing about whether it's safe to vote for one's favorite, LNH specifies that supporting other candidates in addition to a voter's favorite cannot hurt their first choice. This allows candidates to encourage supporters to rank others at no risk to themselves.<sup>6</sup>

Both criteria are clearly desirable, but no deterministic voting method proposed to date has been able to satisfy both for elections with more than two candidates. We posit that rather than passing one but then failing the other criterion badly, voting methods should instead seek to maximize both. We believe that violations of LNH and FB, and their impacts on strategic voting incentives, should be evaluated statistically rather than with an axiomatic approach alone.

As we will show, STAR Voting incentivizes both honest and expressive voting by counting all ballot data given. The scoring round incentivizes voters to give their favorite(s) five stars. The runoff incentivizes voters to also give intermediate scores because showing honest preference order ensures their full vote will go to the finalist they prefer in the automatic runoff. Compare this with Instant Runoff Voting, which can actually incentivize Favorite Betrayal because, in order to pass Later No Harm, it ignores down-ballot voter preferences which could have been relevant.

Despite the widespread claim that IRV eliminates the Spoiler Effect, Emily Dempsey (2018) demonstrates that in order to pass Later No Harm, a voting method must by definition fail to eliminate the Spoiler Effect and vote-splitting. For these reasons, we believe that the adherence to Later No Harm as a desirable pass/fail criterion is problematic.

In IRV, some voters whose favorites are eliminated will have their next choice counted, but voters whose favorites are eliminated in the final round will not. This biases elections against voters who prefer strong underdog candidates with broad support. Counting the full ballot for some voters while ignoring relevant ballot data for others (as Later No Harm requires) gives voters a false sense of agency, may erode trust in the system and in voting reform in general, and is out of keeping with the spirit of one person, one vote.

Elections spoiled due to vote-splitting not only fail to elect the right winner, they also bias outcomes in predictable ways. "The Spoiler Effect occurs when a third candidate entering a race splits votes with a similar candidate who would otherwise win, thus causing a candidate less-preferred by the electorate to win instead." Dempsey (2018)

The center-squeeze Spoiler Effect in particular is pervasive in Plurality but is exhibited by IRV as well. When it happens, it fuels polarization and entrenches two party domination by preventing candidates in the middle of the field from winning. The center-squeeze effect was clearly demonstrated in simulated elections visualized by Ka-Ping Yee (2005) and was analyzed and discussed by Warren Smith and later Mark Frohnmayer (2017). Yee Diagrams such as those shown in Fig. 2 demonstrate that STAR, Score + Top Two Runoff, and Condorcet methods are more accurate than

<sup>6</sup> Later No Harm is defined as "Adding a later preference to a ballot should not harm any candidate already listed. See Appendix A." Woodall (1997).

Yee Diagrams such as these stills from Mark Frohnmayer's "Animated Voting Methods" video can be used to visualize voting method pathologies such as spoiled elections, and to identify center squeeze and expansion effects.

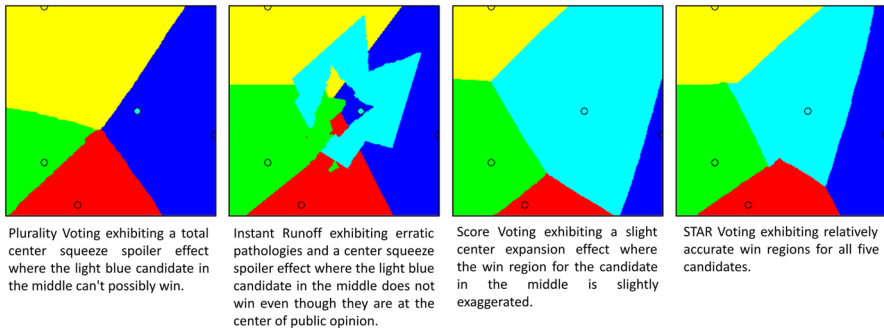


Fig. 2 Yee diagrams

other alternative voting methods and do not exhibit exaggerated center-squeeze or center-expansion biases, two common pathologies which can result in unrepresentative outcomes.

Many systemic problems in electoral politics in general can be traced back to vote-splitting and the Spoiler Effect under Plurality Voting. Fear of wasting one's vote leads directly to "lesser-evil" voting where voters are unable to safely support their favorite. This, in turn, is the reason for the "electability" bias, and it's likely a fundamental reason why political leadership in the United States remains starkly out of sync with the makeup of the voting population.

In research on the demographics of elected officials, (Kauzlarich, 2019) white men held 62% of U.S. elected offices in 2019, despite comprising only 30 percent of the population. In order to achieve gender parity and racial equity in politics—or to overcome two-party domination—we need a level playing field where candidates from underrepresented communities can fairly compete, but the reality is that voting methods biased towards those who are deemed most electable (Abramowitz, 1989) are likely to maintain serious disparities in representation, regardless of public opinion. Fear of vote-splitting appears to be a powerful driver of voter behavior; if so, this would lead to a significant additional advantage for the candidates deemed most electable. In practice, those who benefit from the "electability bias" are usually those who raised the most money, (OpenSecrets, 2020) candidates with name recognition, and incumbents (OpenSecrets, 2019) who tend to be wealthy, older, white, and male.

We posit that by passing the Equality Criterion, vote-splitting caused by the voting method itself can be eliminated. The Equality Criterion states that for any given vote, there is a possible opposite vote, such that if both were cast, it would not change the outcome of an election.<sup>7</sup> The Equality Criterion ensures that if one party had the support of 51% of the voters and ran multiple candidates, and another party had the support of 49% of the electorate and ran only one candidate, the majority

<sup>7</sup> For a rigorous definition of the Equality Criterion, see Appendix A.

faction would always have some way to give all of their candidates full support and thus guarantee a win, even if the front-runners were unknown.

In 1964, *Wesberry v. Sanders*, (Black, 1964) The U.S. Supreme Court declared that equality of voting—one person, one vote—means that “the weight and worth of the citizens’ votes as nearly as is practicable must be the same.” Passing the Equality Criterion ensures that it’s possible for voters who disagree to cast equally weighted and opposite votes, no matter how many candidates are on their side. Approval, Score, Smith/Minimax, and STAR Voting all pass this basic and ‘practicable’ criteria; Plurality and Instant Runoff Voting do not.

#### 4 A non-binary statistical approach

In the axiomatic approach to evaluating voting methods, one begins by clearly defining an outcome that is arguably pathological, and then proceeds to check if that pathology is possible. In the statistical approach, the question is not just “can this pathology ever happen?” but “how problematic is this pathology on average?” Thus, we identify the context where the problem occurs, measure its frequency, and then quantify the severity of its impact.

In principle, one could define a context using only real historical outcomes, but in practice that would mean sample sizes would be unacceptably small, and for newly adopted voting methods there would be no samples at all. Historical results almost never include the data required to assess what role, if any, strategic voting played in an election. In practice, measuring the frequency of various pathologies adequately means defining a generative probability model and using it to run simulated elections.

We simulate elections for a given voting method using the following steps, which each may involve some randomness:

1. *Populate election*: Generate voters and candidates with the inherent characteristics that will be needed for future steps.
2. *Assign utilities*: Decide how each voter feels about each candidate.
3. *Polling*: Conduct some kind of “pre-election polls” of the electorate.
4. *Honest-Naive Voting*: Cast ballots for each voter using their inherent feelings about the candidates. Determine the winner.
5. *Viability-Aware Voting*: Cast ballots for each voter using their inherent feelings about the candidates, available “polling” information about candidate viability, and their understanding of the voting method. Determine the winner.
6. *Targeted Strategic Voting*: For each targeted strategy tested, identify which voters are likely to employ the strategy in question, which ballots to change, and how to change those ballots. Determine the new winner.

Creating a model as outlined above requires balancing realism and simplicity. Models should have a relatively high likelihood of reproducing historical elections, and counterfactual examples should be plausible. At the same time, the model should be relatively easy to describe and reproduce, while avoiding unnecessary complexities which researchers could use to bias results.

Although this paper focuses on STAR Voting, we have done our best to avoid creating a model biased in favor of STAR. Jameson Quinn’s early work on the models presented here, first made public in Quinn (2017), preceded (and was the basis for) serious STAR activism. Quinn had hoped and expected to instead find support for Bucklin-style voting methods such as Majority Judgment (Balinski & Laraki, 2010).

#### 4.1 Voter model

Populating the election and assigning utilities to the voters and candidates is done according to a voter model. In order to select our voter model we first considered the “impartial culture.” (See: Black (1958) for introduction of the idea; Klahr (1966) for its extension to weak orderings; Fishburn and Gehrlein (1980); and Smith (2000) for generative versions and utilitarian interpretations.) We then considered “normally-distributed spatial models” (see: Downs (1957) for the introduction of the idea as a descriptive model; Smith (2000), Tideman and Plassmann (2008), and Green-Armytage et al. (2016) for simulations using the model).

Tsetlin et al. (2003) points out that impartial culture has too many Condorcet cycles. Tideman (2020) agrees and argues that for this and other reasons spatial models are better, but normal spatial models go too far in the other direction, with too few Condorcet cycles.<sup>8</sup> For these reasons we use a clustered spatial model. Because this is non-parametric (that is, because it can include an unbounded number of clusters), it can reproduce real-world election scenarios—including Condorcet cycles—as precisely as desired.

In the clustered spatial model (as in the spatial model), voters and candidates are characterized by their ideal points in a vector space; in the clustered model they are distributed in that space using a common hierarchical Dirichlet structure of Gaussian clusters, similar to a CrossCat model (Mansinghka et al., 2016). Step-by-step we:

1. Assign relative weights to issue dimensions using a stick-breaking Dirichlet process, adding dimensions until the remaining weight falls below a given threshold. This allows differing distributions of issue weights, while ensuring that, on average, issue weight decays exponentially and thus only a finite number of dimensions need to be modeled.
2. Cluster the dimensions themselves into “views” using a Chinese Restaurant Dirichlet process.
3. Within each view, independently cluster the voters using a Chinese Restaurant Dirichlet process. For each voter cluster, assign a mean and variance for each dimension, and draw the voters’ ideal points as normally-distributed using their cluster mean and variance in each view.

<sup>8</sup> Normal distributions are symmetric around a center point along any axis; thus, they tend not to have Condorcet cycles (Davis et al., 1972).



This results in a model with good exchangeability properties that make it relatively easy to sample from and analyze. We believe that a nonparametric model of this form will show good realism, and in particular that specific kinds of potentially-pathological scenarios such as Condorcet cycles will occur with a realistic frequency – neither artificially often as in impartial culture, nor artificially rarely as in normally-distributed spatial models.

## 4.2 Polling model

Next we define the polling and media models used. “An initial “poll” is performed using Approval Voting, where the approval threshold for each voter is set at 70% of the way between an average utility candidate (0%) and the voter’s favorite (100%). We use Approval Voting for the poll because it gives the voters a realistic amount of information about the relative strength of the candidates. We simulate real-world polling error with Gaussian noise (standard deviation of 5%).

## 4.3 Initial strategy models

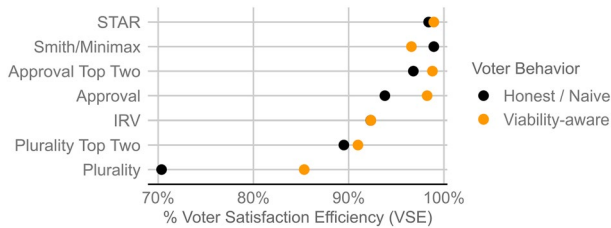
We consider both a zero-information “honest/naive” strategy and a “viability-aware” strategy using the Approval polling discussed above.<sup>9</sup>

For Approval methods, the approval threshold for the naive strategy is set at 40% of the way up from the average candidate to the voter’s favorite. For STAR, the “naive” strategy is to vote per the instructions by normalizing one’s utilities to the [0–5] star scale and voting accordingly.

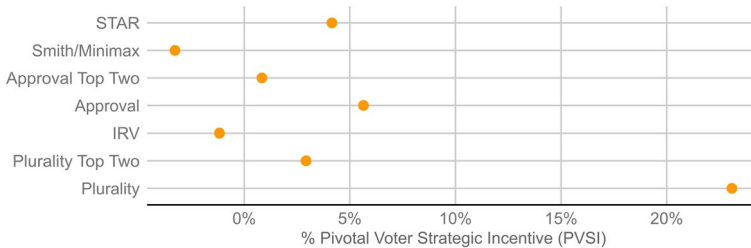
“Viability-Aware” strategies are as follows:

- *Plurality*: For a given voter, determine the expected utility of the election ( $EV$ ), then, given a candidate with utility  $u_i$  and estimated probability of winning  $p_i$ , score candidates equal to  $p_i(u_i - EV)$ . Vote for the candidate with the highest score.
- *Plurality Top Two*: Same as Plurality except that voters who do not prefer the polling leader replace the estimated probability of winning with an estimate of the probability of being in a two-way tie for second.
- *Approval*: Determine the expected utility of the election, and vote for every candidate with utility greater than that.
- *Approval Top Two*: Same as Approval, but replace the estimated probability of winning with the estimated probability of being in a two-way tie for second place.

<sup>9</sup> Specifically, each candidate’s probability of winning is estimated by assuming that their true percent approval is independently beta-distributed around their observed approval, tuning the total  $\alpha + \beta$  of the two beta distribution parameters to have a given margin of error if  $\alpha = \beta$ .



**Fig. 3** Voter Satisfaction Efficiency of selected voting methods. Model for Fig. 3 is run using 10,000 electorates each with 6 candidates and 5001 voters. Model for Figs. 4 and 5 are run using 10,000 electorates each with 6 candidates and 101 voters. Insofar as voters are naturally clustered, the simulated voters stand in for a greater number of actual individuals



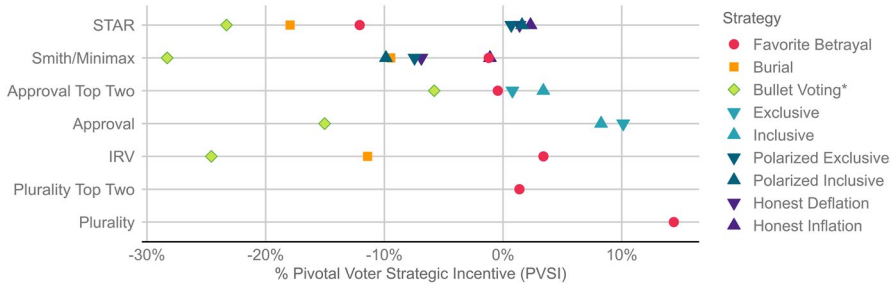
**Fig. 4** A “viability-aware” strategy is incentivized over honest/naive behavior for most methods tested but the size of these incentives varies considerably. Honest voting outperformed our viability-aware strategy for Smith/Minimax and IRV

- *Smith/Minimax*: Use the same candidate scores determined for Plurality and rank the candidates in decreasing order by score. This can result in both the burial of disliked frontrunners and favorite betrayal akin to what is done in plurality.
- *IRV*: Same as Smith/Minimax, but rank the candidates with negative scores honestly. This can result in favorite betrayal but not burial.
- *STAR*: Balance strategic exaggeration (using 5’s and 0’s instead of 3’s and 2’s to maximize influence in the scoring round) with the competing incentive of not wanting to give multiple viable candidates the same score. (This yields more 5’s and 0’s than the naive strategy and is mostly unaffected by the presence of non-viable candidates in the race.)

Note that the above viability-aware strategies exhibit honest preferences for Approval, Approval Top Two and STAR, but may exhibit dishonest preference order for Plurality, Plurality Top Two, IRV, and Smith/Minimax. This is due to the fact that there’s no way to change an ordinal ballot that wouldn’t also change the preference order, whereas in cardinal methods there is more flexibility.

## 5 Representative outcomes: Voter Satisfaction Efficiency

An ideal winner represents as many voters as possible as well as possible. One way to measure this is with utility. Voter Satisfaction Efficiency (VSE) is a linear measure of a voting method’s utilitarian outcomes. In VSE a voting method that always



**Fig. 5** Incentives for various targeted strategies compared to the most-incentivized behavior from Fig. 4 (honest/naive or viability-aware). \*Bullet voting, unlike other strategies pictured, is employed by a random faction. This is because any faction could realistically bullet vote, so targeting would have required potentially unrealistic assumptions

elected this ideal candidate would score 100%, while one that elected a candidate completely at random would score 0%.

Smith/Minimax and STAR Voting delivered the highest levels of Voter Satisfaction of all methods tested (see Fig. 3). Viability-aware voting yielded higher VSE under all voting methods where it was strategically incentivized (see Fig. 4). Approval-based methods also performed notably well under best-case-scenario voter behavior.<sup>10</sup>

## 6 Honest and strategic incentives

A voting method’s outcomes can only be as good as the ballot data itself, and ballot data is likely to only be as good as the voter behavior. In order to evaluate the likelihood of specific voter behaviors and their impacts on the quality of electoral outcomes, we define a statistical metric called Pivotal Voter Strategic Incentive (PVSI). This measures strategic incentives by taking a faction who wants to employ a specific strategy and changing their votes, starting with the person who stands to gain the most and continuing down the list, until the outcome changes with one pivotal voter. We then evaluate the change in outcome and resulting strategic incentives from this voter’s perspective.<sup>11</sup> When a strategy works as intended its incentive will be greater if it takes relatively low coordination, it will average near zero if the strategy rarely has any effect, and it will be negative if the risk of the strategy backfiring outweighs the potential reward of success.

<sup>10</sup> See Appendix B for more VSE results, varying the model, candidates, and parameters. See supplementary material for VSE/PVSI code in Python.

<sup>11</sup> PVSI is measured as a percentage and is defined as the average over all simulations of the pivotal voter’s utility of the winner under strategic voting, minus the utility of the candidate who wins in the absence of strategy, divided by the voter’s utility gap between their favorite and their average candidates. Simulation runs in which the outcome never changes count as the pivotal voter receiving no change in utility.

We begin by using Pivotal Voter Strategic Incentive to look at how much this voter benefits if a random faction<sup>12</sup> that includes them takes candidate viability into consideration (Viability-Aware).

We then consider the impacts of strategy types such as “Favorite Betrayal” (rating the “allied” frontrunner above your true favorite); “Burial” (rating the “enemy” frontrunner below apparently non-viable candidates you actually like less); “Bullet Voting\*” (voting for only your favorite); “Inclusive” and “Exclusive” strategies in Approval Voting (lowering or raising your approval threshold); “Polarized Inclusive” and “Polarized Exclusive” in STAR Voting (scoring all candidates 5 or 0); and “Honest Inflation” and “Honest Deflation” in STAR Voting (exaggerating candidates scores to 5, 4, 1, or 0 in order to boost an allied frontrunner or block an enemy frontrunner).

For targeted strategies (all those in Fig. 5 except for Bullet Voting), we identify the faction who stands to benefit most from strategic voting.<sup>13</sup>

In Fig. 5, we can see that in STAR Voting, the dishonest strategies (Favorite Betrayal, Burial, and Bullet Voting) are all strongly disincentivized. Weakly incentivized honest/semi-honest strategies include Polarized Inclusive, in which voters might give all the candidates on their side 5 stars, and Honest Inflation, in which a voter might give their top candidates 5 stars, their compromise candidates 4, an opposing frontrunner 1, and their last choice 0.

In Instant Runoff Voting, the most-incentivized strategy is Favorite Betrayal, commonly known as “voting for the lesser of two evils”, at 3%. The positive incentive here indicates that ranking your favorite 1st can backfire. This is even more true in Plurality, where Favorite Betrayal is incentivized at 14%. (PVSI would be 100% if a tactical vote was guaranteed to change the winner from an average candidate to the pivotal voter’s favorite.)

In Approval, the most incentivized strategy, topping out at 10%, is to set one’s approval threshold between the two most popular candidates. With Approval Top Two, this goes down to 3%. For STAR there is a weak incentive, at 2%, to exaggerate one’s scores up while maintaining an honest preference order. In STAR, Approval Top Two, and Smith/Minimax, strategies that give less than full support to a voter’s favorite are all disincentivized. Smith/Minimax disincentivized strategic voting across the board.

We can clearly see that Plurality Voting, with its strong and transparent strategic incentives, pressures voters to support “electable” candidates, further entrenching barriers to entry and existing inequities in political representation. While voters trying a new method for the first time may need to unlearn strategic habits, the Pivotal Voter Strategic Incentive voter models illustrate why we believe that over time voter behavior would become more honest under STAR and other top-tier voting methods.

<sup>12</sup> A random faction is defined by randomly choosing two candidates and taking those voters who prefer the first of them.

<sup>13</sup> In IRV and Plurality Top Two, voters who prefer the original third-place candidate over the original winner make up the strategic faction. In the other methods it’s voters who prefer the original second-place candidate.

## 7 Conclusion

STAR Voting’s expressive and user-friendly ballot, simple and transparent tabulation, positive incentives to vote one’s conscience, and its accurate outcomes make it an actionable solution to level the playing field and elect representative winners in single-winner elections.

## Appendices

### Appendix A: Formal voting definitions and criteria

Consider a voting method as a set of algorithms, indexed by a finite set of candidates  $\mathcal{C}$ .<sup>14</sup> For each such set, the method defines a space  $\mathcal{B}$  of valid ballots; and for any list of  $V$  ballots  $(b_1, \dots, b_V)$ , where  $b_i \in \mathcal{B}$  for all  $1 \leq i \leq V$ , the method defines a “result”  $R(b_1, \dots, b_V)$ , which in the most general case is a lottery over subsets of  $\mathcal{C}$ . In this paper, we are focused on single-winner voting methods, so  $R$  should be a lottery over individual members of  $\mathcal{C}$ , and in fact the voting methods we consider here are all deterministic except in rare cases of ties (asymptotically measure-zero as number of voters goes to infinity).

Let us call a given list of ballots an “election”, and denote specific elections we consider as  $\mathcal{E}$ ,  $\mathcal{E}^*$ , etc. We also assume that any ballot in  $b \in \mathcal{B}$  is legible at least in terms of what strict preferences it expresses;<sup>15</sup> that is to say, for any two candidates  $X, Y \in \mathcal{C}$ , either  $X > Y$  or  $X \leq Y$  holds, which is to say that ballot  $b$  either does or doesn’t strictly prefer  $X$  over  $Y$ . The expected axioms for strict preference hold:

- $X \stackrel{b}{\leq} Y$ ,
- $X \stackrel{b}{>} Y$  implies  $Y \stackrel{b}{\leq} X$ ,
- $X \stackrel{b}{>} Y$  and  $Y \stackrel{b}{>} Z$  imply  $X \stackrel{b}{>} Z$ .

**Definition 1 (Favorite Betrayal Criterion (FBC))** Consider an election  $\mathcal{E} \equiv \{b_1, \dots, b_V\}$  and its result  $R$ . Take the voter who cast ballot  $b_i$  (for an arbitrary  $i$ ) and assume that they have an arbitrary real-valued utility  $u_{ij}$  for each candidate  $c_j \in \mathcal{C}$ . Denote their favorite candidate,  $\arg \max_j u_{ij}$ , as  $X$ .<sup>16</sup>

In order to pass FBC: For any “strategic” ballot  $b_i^*$  that voter  $i$  could have cast, thus giving a new result  $R(b_1, \dots, b_{i-1}, b_i^*, b_{i+1}, \dots, b_V)$  with all other ballots held constant, there must be some “non-betraying strategic ballot”  $b_i^{**}$  such that the following both hold:

- $Y \stackrel{b_i^{**}}{\leq} X$  for any  $Y$ ,

<sup>14</sup> Unless otherwise noted, we will consider the set of candidates as fixed, and thus suppress  $\mathcal{C}$  as a subscript on other entities such as  $\mathcal{B}$ .

<sup>15</sup> Note, of course, that the assumption that ballots carry strict preference information does not imply that they don’t also carry further information such as strength of preference.

<sup>16</sup> If there are several candidates tied for voter  $i$ ’s favorite, this must hold for all of them.

- Voter  $i$ 's expected (Von Neumann–Morgenstern) utility for  $R(b_1, \dots, b_{i-1}, b_i^*, b_{i+1}, \dots, b_V)$  is not greater than that for  $R(b_1, \dots, b_{i-1}, b_i^{**}, b_{i+1}, \dots, b_V)$ .

◇

For the next criterion, we must impose a bit more structure on the set of valid ballots. For any ballot  $b$  and subset of candidates  $\mathcal{S}$ , there must be a single ballot  $b^{-\mathcal{S}}$  which is the same as  $b$  except that it “removes all support for” the candidates in  $\mathcal{S}$ . The following must hold:

- For  $X, Y \in \mathcal{S}$ :  $X \stackrel{b^{-\mathcal{S}}}{>} Y$
- For  $X \notin \mathcal{S}, Y \in \mathcal{S}$ :  $X \stackrel{b^{-\mathcal{S}}}{>} Y$
- For  $X, Y \notin \mathcal{S}$ :  $(X \stackrel{b}{>} Y \Leftrightarrow X \stackrel{b^{-\mathcal{S}}}{>} Y)$

**Definition 2** (*Later No Harm (LNH)*) Consider an election  $\mathcal{E} \equiv \{b_1, \dots, b_V\}$ . In order to pass LNH, there must not exist a voter  $i$  and a candidate subset  $\mathcal{S}$  such that the following hold:

- For every  $X \notin \mathcal{S}, Y \in \mathcal{S}$ :  $X \stackrel{b_i}{>} Y$ ,
- There is some candidate  $X \notin \mathcal{S}$  who is more likely to win in  $R(b_1, \dots, b_{i-1}, b_i, b_{i+1}, \dots, b_V)$  than in  $R(b_1, \dots, b_{i-1}, b_i^{-\mathcal{S}}, b_{i+1}, \dots, b_V)$ .

◇

Finally, we will define the Equality Criterion:

**Definition 3** (*Equality Criterion*) In order to pass the Equality Criterion, for every ballot  $b$ , there must exist some (not necessarily distinct) ballot  $\tilde{b}$ , such that for any list of ballots  $b_1, b_2, \dots, b_n$ ,

$$R(b_1, b_2, \dots, b_n, b, \tilde{b}) = R(\alpha(b_1, b_2, \dots, b_n))$$

for every permutation  $\alpha$  of  $b_1, b_2, \dots, b_n$ . Note that the equality in the above equation is in distribution.

Additionally,  $\tilde{b}$  must have the opposite preferences of  $b$ . That is,

- $X \stackrel{b}{>} Y \Leftrightarrow Y \stackrel{\tilde{b}}{>} X$ ,
- $X \stackrel{b}{\leq} Y \Leftrightarrow Y \stackrel{\tilde{b}}{\leq} X$ .

◇

The Equality Criterion implies the anonymity Criterion and therefore the identical input options Criterion. It is also easy to show that the Equality Criterion combined with the mutual unanimity Criterion (the requirement that if all voters prefer a subset of candidates over all candidates outside that subset, a member of the subset will win) implies the weak mutual majority Criterion (the requirement that if a majority of voters all prefer a subset over all members outside that subset, there is a way for them to vote to ensure that a member of the subset will win).

## Appendix B: VSE and PVS: robustness to model and parameters

In order to test the robustness of VSE and PVS, we have run election simulations with changes to models and parameters, showing that results remain broadly similar across such changes.

The following graphs are each made using 4000 simulated elections. Unless otherwise specified, they use the following parameters, like the figures in the main paper:

- *textitModel*: Hierarchical Dirichlet (CrossCat-like), as described in the main paper. Dirichlet alpha (clustering) parameters: dimensional clustering 1, dimensional weights within a cluster 1, voter clusters 1.
- *Candidates*: 6
- *Voters*: 101
- *“Pickiness” in initial Approval poll*: 0.7 (This means that in the initial approval poll, voter  $i$  will “support” candidate  $j$  if and only if  $U(i, j) > (Max_k[U(i, k)] - Min_k[U(i, k)]) * 0.7 + Min_k[U(i, k)]$ , where  $U(i, j)$  signifies the utility of voter  $i$  for candidate  $j$ .)
- *Standard deviation of polling noise (both for initial approval poll, and for using viability-aware results to define strategy)*: 5% of electorate. (This is like a real-world poll with a margin of error, including both sampling error and systematic error, of 10%. Roughly speaking, this would typically indicate a sampling-only error of around 5%.)

As in the main paper, Bullet Voting is not included in strategy graphs for Plurality or Plurality Top Two because it is non-applicable. Plurality ballots do not allow any non-bullet votes.

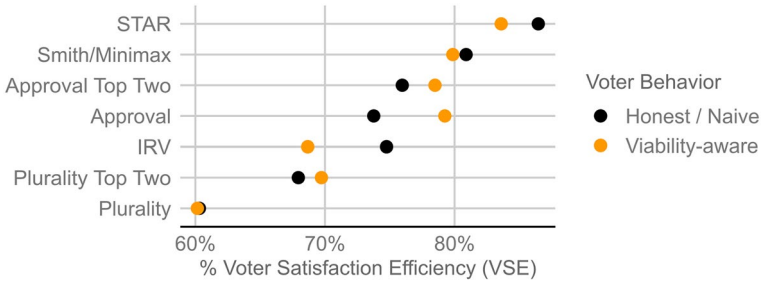
Favorite Betrayal is also omitted from the strategy graphs for Approval voting in the main paper and appendices. This is because the authors were concerned that without further explanation, this data point could be easily misunderstood to incorrectly imply that voters should not approve their favorite. On the contrary, Approval passes Favorite Betrayal Criterion and voters should always approve their favorite candidate(s).

In Approval Voting (with no Top Two runoff), the most incentivized strategy, at 10.1%,<sup>17</sup> is “Exclusive”, meaning that a voter would likely do best to approve their preferred frontrunner and all candidates they prefer to that candidate, including their favorite, but no others. That said, Favorite Betrayal was less strongly incentivized for Approval at 9.8%. This positive incentive indicates that if voters do betray their favorite in Approval, it is unlikely to backfire and hurt the voter, or to hurt the election outcome, especially if their assessments for candidate viability are accurate.

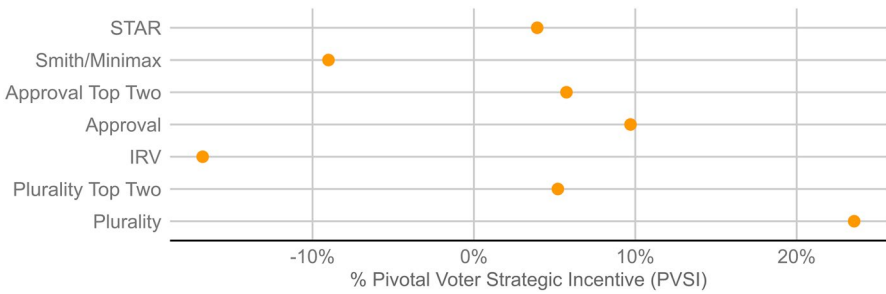
These incentives, and their potential impacts, are important to understand because Approval Voting has an identical-looking ballot to Plurality, so it’s to be expected that some voters used to voting in Plurality might continue to betray their favorite even if it’s no longer necessary.

<sup>17</sup> Strategic Incentive numbers given here for Approval Voting are as they would appear in Fig. 5 in the main paper.

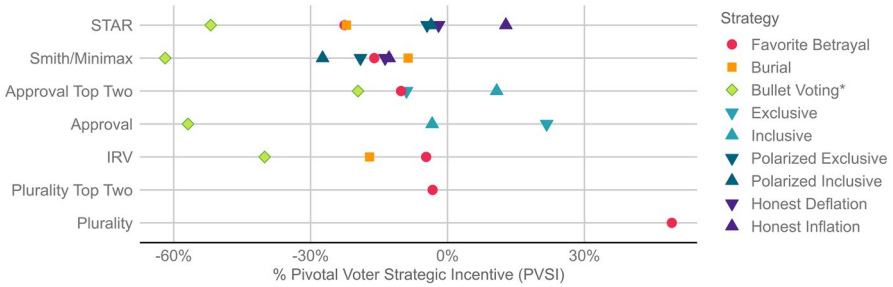
– **Model:** Impartial culture (Figs. 6, 7, 8).



**Fig. 6** VSE results. 6 candidates. Impartial Culture model



**Fig. 7** PVSI of viability aware vs. honesty. 6 candidates. Impartial culture model



**Fig. 8** PVSI of strategies over viability-awareness. 6 Candidates. Impartial culture model



- **Model:** 2 equally-weighted ideology dimensions; independent Gaussian distributions; utilities equal to negative of Euclidean distance (Figs. 9, 10, 11).

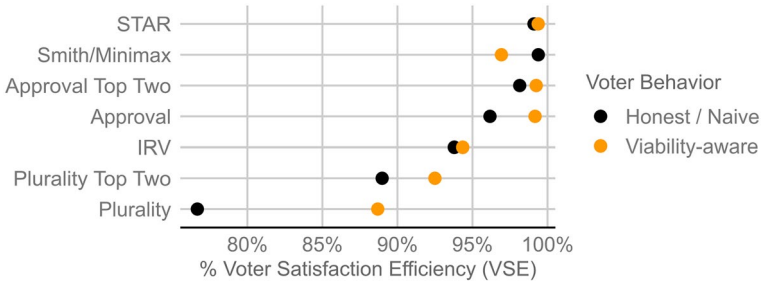


Fig. 9 VSE results. 6 candidates. 2-D spatial model

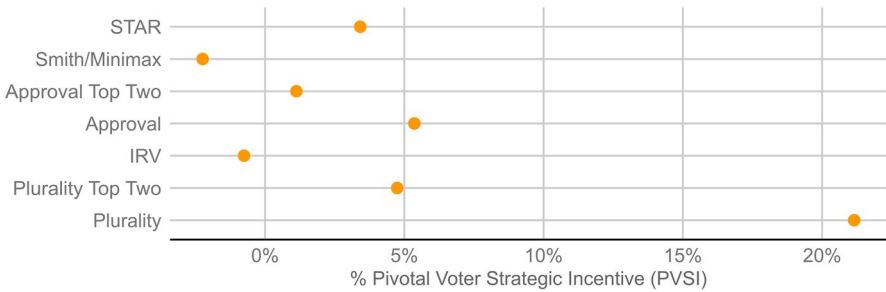


Fig. 10 PVSI of viability-awareness vs. honesty. 6 candidates. 2-D spatial model

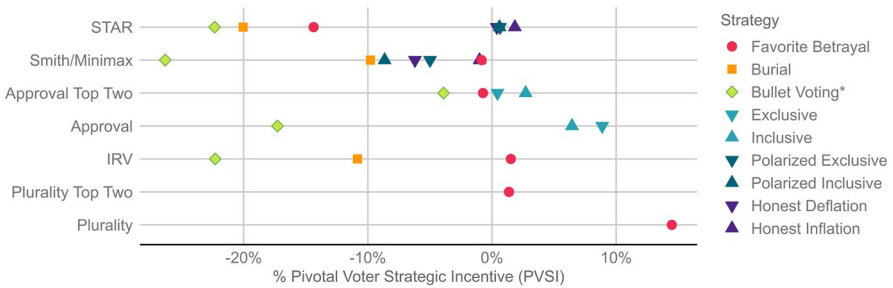


Fig. 11 PVSI of strategies vs. viability-awareness. 6 candidates. 2-D spatial model

- **Model:** 3 equally-weighted ideology dimensions; independent Gaussian distributions; utilities equal to negative of Euclidean distance (Figs. 12, 13, 14).

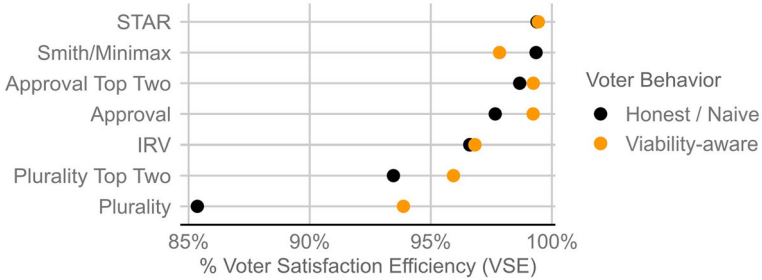


Fig. 12 VSE results. 6 Candidates. 3-D spatial model

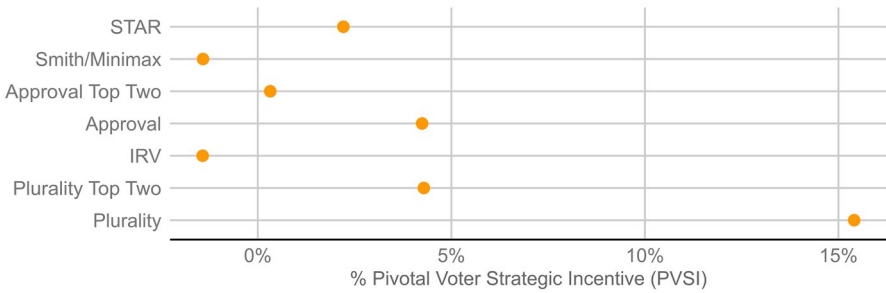


Fig. 13 PVSIs of viability-awareness vs. honesty. 6 candidates. 3-D spatial model

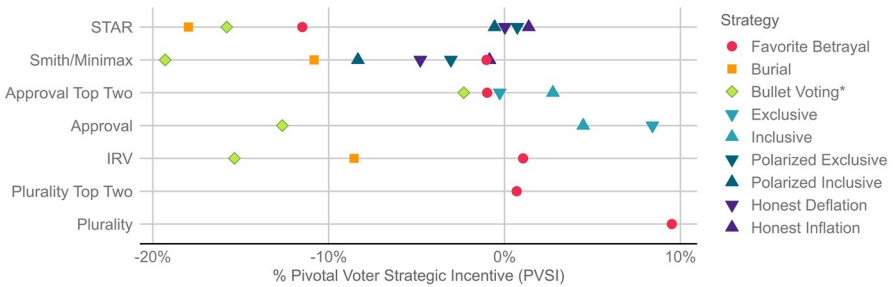
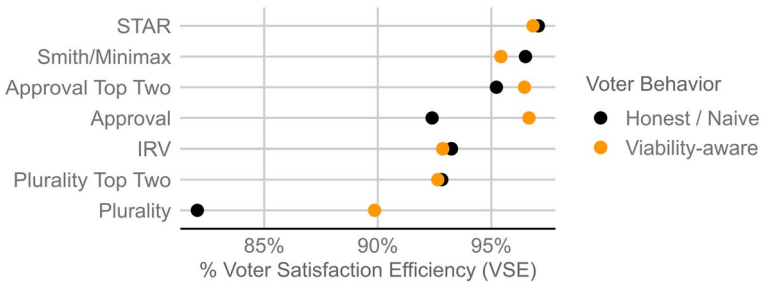
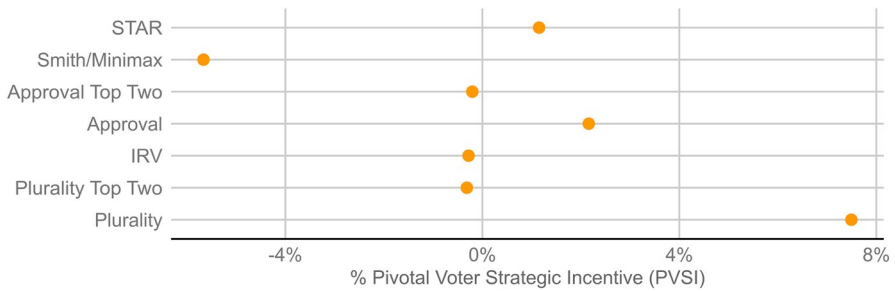


Fig. 14 PVSIs of strategies vs. viability-awareness. 6 candidates. 3-D spatial model

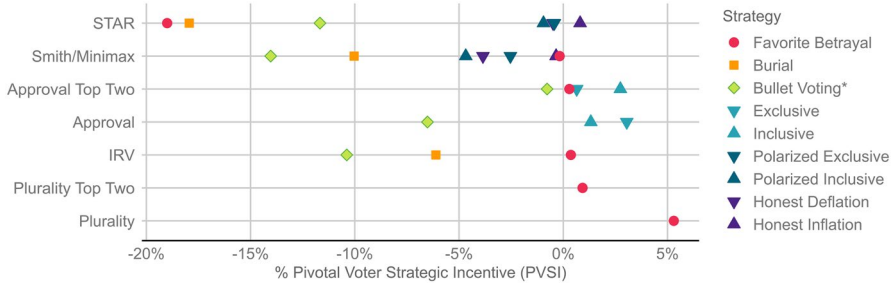
– **Candidates: 3** (Figs. 15, 16, 17)



**Fig. 15** VSE results. 3 candidates. Hierarchical Dirichlet model

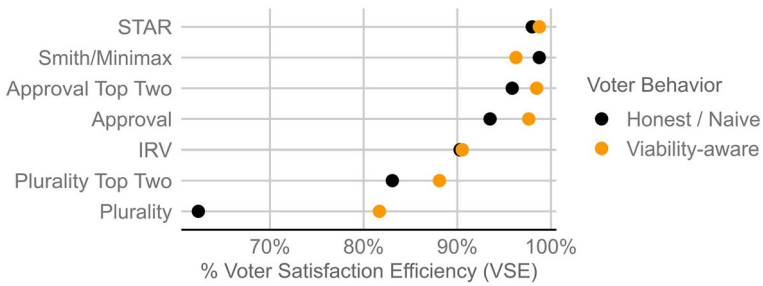


**Fig. 16** PVSI of viability-awareness vs. honesty. 3 cand. Hierarchical Dirichlet model

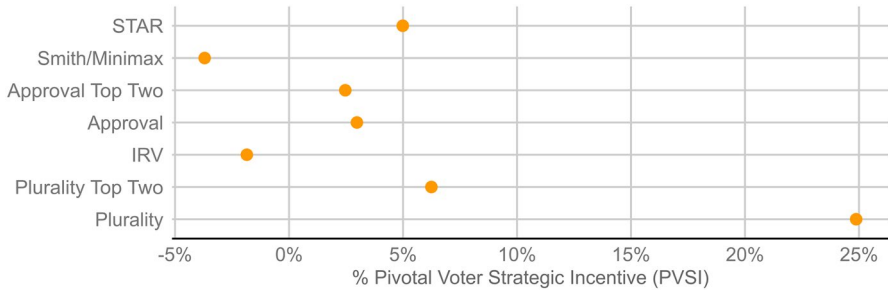


**Fig. 17** PVSI of strategies vs. viability-awareness. 3 cand. Hierarchical Dirichlet model

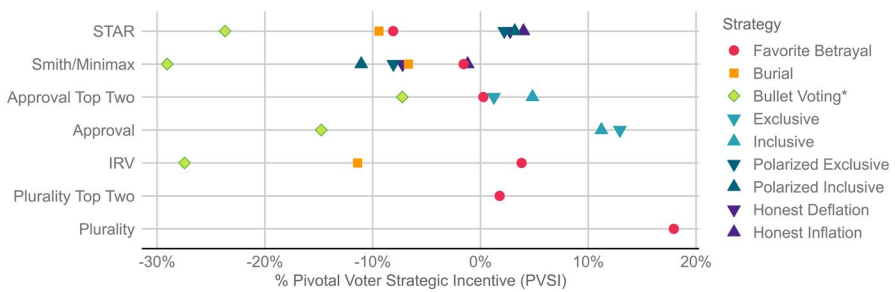
– **Candidates:** 9 (Figs. 18, 19, 20).



**Fig. 18** VSE results. 9 candidates. Hierarchical Dirichlet model

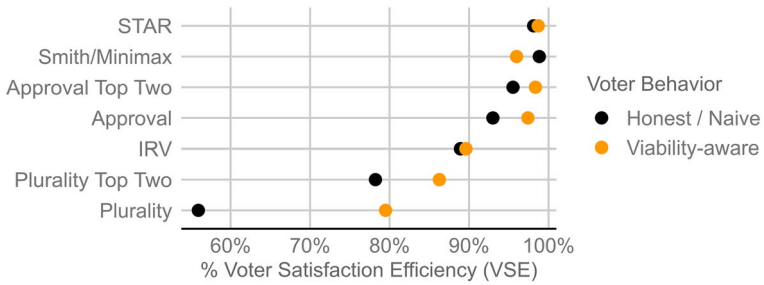


**Fig. 19** PVSIs of viability-awareness vs. honesty. 9 cand. Hierarchical Dirichlet model

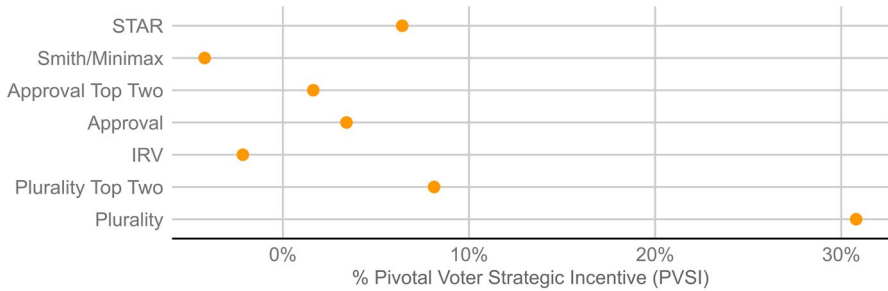


**Fig. 20** PVSIs of strategies vs. viability-awareness. 9 cand. Hierarchical Dirichlet model

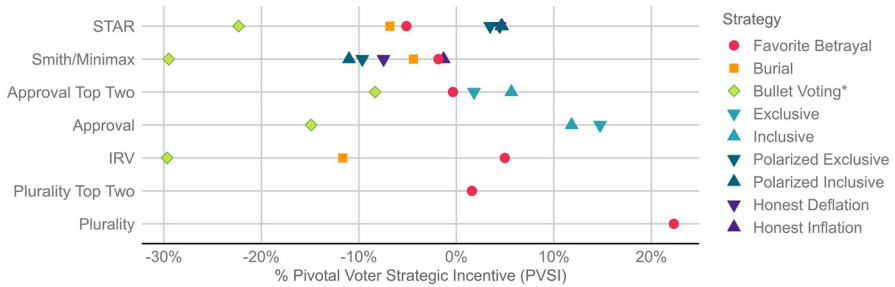
– **Candidates:** 12 (Figs. 21, 22, 23).



**Fig. 21** VSE results. 12 candidates. Hierarchical Dirichlet model



**Fig. 22** PVSIs of viability-awareness vs. honesty. 12 cand. Hierarchical Dirichlet model



**Fig. 23** PVSIs of strategies vs. viability-awareness. 12 cand. Hierarchical Dirichlet model

- **“Pickiness” in initial Approval poll: 99%** (This means that voters will almost certainly bullet vote in the initial approval poll.) (Figs. 24, 25, 26).

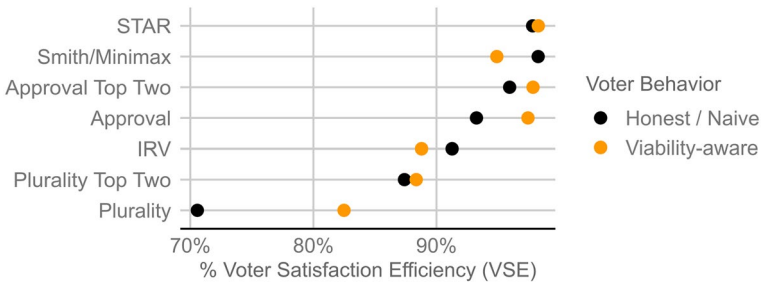


Fig. 24 VSE results with “pickiness” 99%. 6 candidates. Hierarchical Dirichlet model

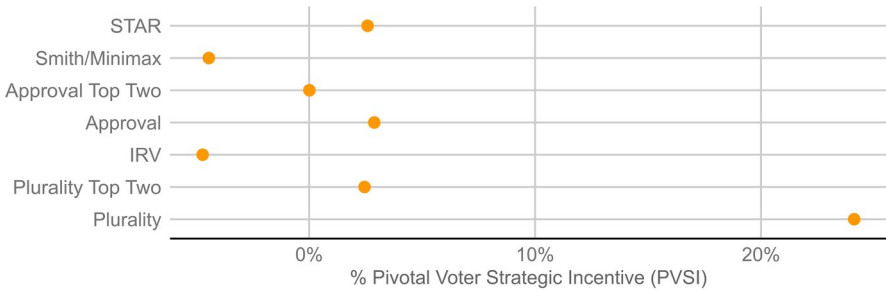


Fig. 25 PVSIs of viability-awareness vs. honesty with pickiness 99%. 6 cand. Hier. Dir. model

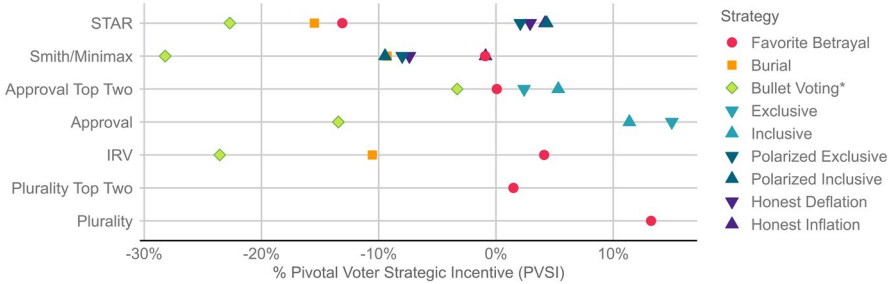


Fig. 26 PVSIs of strategies vs. viability-awareness with pickiness 99%. 6 cand. Hier. Dirichlet

- **Standard deviation of polling noise: 0%**. (Perfectly-accurate polling) (Figs. 27, 28, 29)

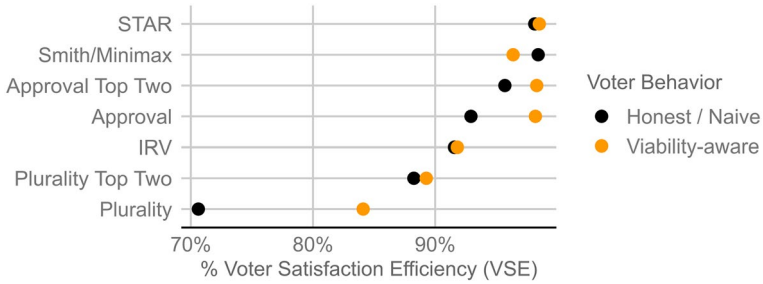


Fig. 27 VSE results with accurate polling. 6 candidates. Hierarchical Dirichlet model

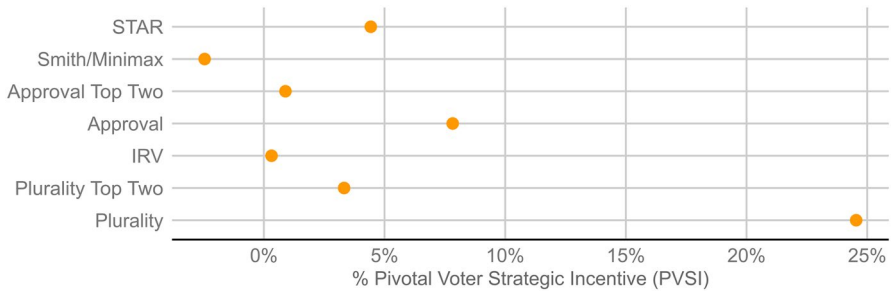


Fig. 28 PVSI of viability-awareness vs. honesty with accurate polling. 6 candidates. Hierarchical Dirichlet model

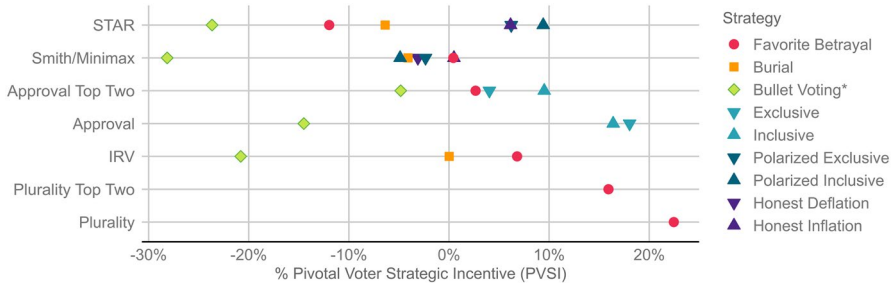
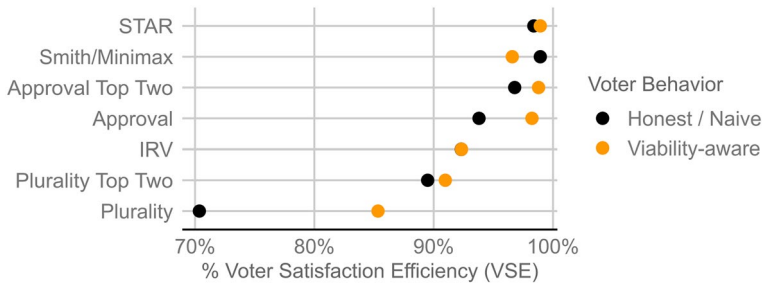
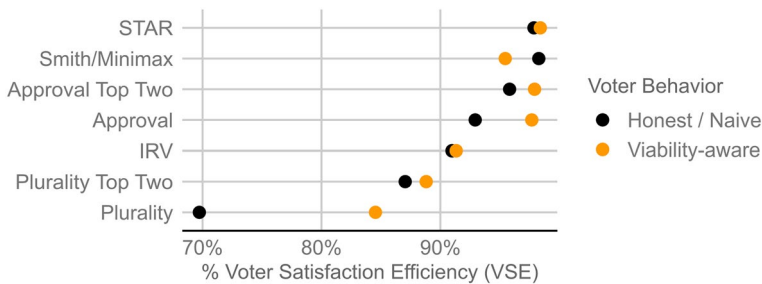


Fig. 29 PVSI of strategies vs. viability-awareness with accurate polling. 6 candidates. Hierarchical Dirichlet model

– **Voters:** 101 vs. 5001 (Figs. 30, 31)



**Fig. 30** VSE results. 101 voters. 6 candidates. Hierarchical Dirichlet model



**Fig. 31** VSE results. 5001 voters. 6 candidates. Hierarchical Dirichlet model

When compared side by side, the above two figures illustrate the minimal differences between a voter model with 101 voters and a voter model with 5001 voters.

**Acknowledgements** Special thanks to Mark Frohnmayer for contributions as STAR Voting co-inventor and advocacy for the equally weighted vote.

## Declarations

**Conflict of interest** Sara Wolk is the Executive Director for the Equal Vote Coalition, which advocates for some of the voting methods discussed here.

## References

- Abramowitz, A. I. (1989). Viability, electability, and candidate choice in a presidential primary election: A test of competing models. *The Journal of Politics*, 51(4), 977–992. <https://doi.org/10.2307/2131544>.
- Arrow, K. J. (1950). A difficulty in the concept of social welfare. *Journal of Political Economy*, 58(4), 328–346.



- Balinski, M., & Laraki, R. (2010). *Majority Judgment*. Cambridge: The MIT Press.
- Black, D. (1958). *The theory of committees and elections*. Cambridge: Cambridge University Press. <https://doi.org/10.1007/978-94-009-4225-7>.
- Black, H. (1964). *Wesberry v. Sanders*, 376 U.S. 1.
- Davis, O. A., DeGroot, M. H., & Hinich, M. J. (1972). Social preference orderings and majority rule. *Econometrica*, 40(1), 147–157. <https://doi.org/10.2307/1909727>.
- Dempsey, E. (2018). Farewell to Pass/Fail. [https://starvoting.org/pass\\_fail](https://starvoting.org/pass_fail)
- Downs, A. (1957). An economic theory of political action in a democracy. *Journal of Political Economy*, 65(2), 135–150.
- Fishburn, P. C., & Gehrlein, W. V. (1980). The paradox of voting: Effects of individual indifference and intransitivity. *Journal of Public Economics*, 14(1), 83–94. [https://doi.org/10.1016/0047-2727\(80\)90006-7](https://doi.org/10.1016/0047-2727(80)90006-7)
- Frohnmayr, M. (2017). Animated Voting Methods. <https://youtube.com/watch?v=-4FXLQoLDBA>
- Green-Armytage, J., Tideman, T. N., & Cosman, R. (2016). Statistical evaluation of voting rules. *Social Choice and Welfare*, 46(1), 183–212. <https://doi.org/10.1007/s00355-015-0909-0>
- Kauzlarich, N. (2019). *The Electability Myth: The Shifting Demographics of Political Power in America*. Technical report, Reflective Democracy Campaign.
- Klahr, D. (1966). A computer simulation of the paradox of voting. *The American Political Science Review*. <https://doi.org/10.2307/1953365>.
- Maine Supreme Judicial Court. (2017). Opinion of the justices of the supreme judicial court under the provisions of Article VI, Section 3 of the Maine Constitution (oj-17-1).
- Mansinghka, V., Shafto, P., Jonas, E., Petschulat, C., Gasner, M., Tenenbaum, J. B., et al. (2016). Cross-Cat: A fully Bayesian nonparametric method for analyzing heterogeneous, high dimensional data. *Journal of Machine Learning Research*, 17, 1–49.
- Neely, F., & McDaniel, J. (2015). Overvoting and the equality of voice under instant-runoff voting in San Francisco. *California Journal of Politics and Policy*. <https://doi.org/10.5070/P2cjjp7428929>.
- OpenSecrets. (2019). Reelection rates over the years. <https://opensecrets.org/elections-overview/reelection-rates>
- OpenSecrets. (2020). Did money win? <https://opensecrets.org/elections-overview/winning-vs-spending?cycle=2020>
- Quinn, J. (2017). Voter Satisfaction Efficiency (VSE) FAQ. <https://electionscience.github.io/vse-sim/VSE/>
- Simms, L., Zelazny, K., Williams, T., & Bernstein, L. (2019). Does the number of response options matter? Psychometric perspectives using personality questionnaire data. *Psychological Assessment*, 31(4), 557–566. <https://doi.org/10.1037/pas0000648>.
- Smith, W. D. (2000). Range voting. <https://web.archive.org/web/20210801134240/https://rangevoting.org/WarrenSmithPages/homepage/rangevote.pdf>
- The Equal Vote coalition. (2020). 2020 DNC Presidential Delegate Elections. [https://starvoting.org/2020\\_dnc\\_delegate\\_selection](https://starvoting.org/2020_dnc_delegate_selection)
- Tideman, N. (2010). Structure of the election-generating universe. In *Voting power in practice*, Normandy.
- Tideman, T., & Plassmann, F. (2008). *Evaluating voting methods by their probability of success: An empirical analysis*. Virginia Polytechnic Institute and State University
- Tsetlin, I., Regenwetter, M., & Grofman, B. (2003). The impartial culture maximizes the probability of majority cycles. *Social Choice and Welfare*, 21(3), 387–398.
- Woodall, D. R. (1997). Monotonicity of single-seat preferential election rules. *Discrete Applied Mathematics*, 77(1), 81–98.
- Yee, K. P. (2005). *Voting simulation visualizations*. <https://zesty.ca/voting/sim/>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.