Comparing strategic voting incentives in plurality and instant-runoff elections[1]

Andrew C. Eggers[†]

Tobias Nowacki[†]

Abstract

Reformers and researchers often speculate that some voting systems induce less strategic voting than others, but existing research is unhelpful in assessing these conjectures because it is overwhelmingly based on the unrealistic assumption that voters know exactly how others will vote. We propose a general approach to assessing strategic voting incentives given realistic uncertainty about election outcomes. We use this approach to compare strategic voting incentives in three-candidate plurality and instant-runoff (IRV) elections, drawing on preference data from 160 electoral surveys. We show that the common conjecture that strategic voting incentives are smaller in IRV than in plurality is correct: for the full range of beliefs about how strategic one thinks other voters will be, the expected benefit of being a strategic voter rather than a sincere voter is smaller in IRV than in plurality.

[1]This version: May 4, 2019.

[†]Nuffield College and Department of Politics and International Relations, University of Oxford, UK. email: `andrew.eggers@nuffield.ox.ac.uk`

[†]Stanford University. email: `tnowacki@stanford.edu`

1 Introduction

Social choice theory tells us that no reasonable voting system is completely immune from strategic voting (cites). It is reasonable to suspect, however, that some systems are likely to be more susceptible to strategic voting than others. For example, in advance of the 2011 UK referendum to replace plurality parliamentary elections with a form of instant-runoff voting (known as the alternative vote in the UK), Deputy Prime Minister Nick Clegg claimed that IRV "stops people from voting tactically and second-guessing how everybody else will vote in their area";[2] *other non-academic claims here.* Prominent scholars have made similar conjectures. Cox (1997, 95), for example, notes that more information is needed to vote strategically in IRV than in plurality, as does Renwick (2011, pp. 6-7), who concludes that IRV would "reduce but not eliminate incentives for tactical voting" compared to plurality, while Dummett (1984, p. 228) maintains that "a voter who has understood the workings of the procedure, and who has some information about the probable intentions of the others, will have nearly as much incentive to vote strategically" in IRV as in plurality elections.

Considering that the question of which system is more susceptible to strategic voting seems to be relevant to many policymakers and academic researchers alike, previous research provides a surprisingly unsatisfying answer. As noted above, the classic theorems show that all reasonable systems could produce a *manipulable* voting result, i.e. a configuration of ballots such that one or more voters would be better off submitting an insincere ballot than a sincere one. Building on these theorems, a large literature (mostly in mathematics and computer science) has assessed the likelihood of such a manipulable result for different voting rules given some assumption about the distribution of possible voting outcomes (e.g. Chamberlin, 1985; Nitzan, 1985; Saari, 1990; Favardin and Lepelley, 2006).[3] But these studies are of little use for assessing susceptibility to strategic voting, even if we accept their assumptions about likely voting outcomes, because they do not take into account uncertainty about others' votes: they essentially ask how likely a voter is to *regret* a sincere vote after the election takes place, not how likely a voter is to *foresee*

[2] "The Coalition Government's programme of political and constitutional reform: Oral and written evidence", 15 July 2010, HC 358-i, published 22 October 2010 (link). By "tactical voting" Clegg likely means submitting a ballot that differs from one's sincere preference.

[3] A closely related set of papers assesses the probability of results that violate monotonicity and other desirable properties of choice rules (Plassmann and Tideman, 2014; Ornstein and Norman, 2014; Miller, 2017).

that an insincere vote would be optimal before the election takes place. A system can only be susceptible to strategic voting if voters could anticipate, given reasonable beliefs about the relative likelihood of different election outcomes, that an insincere vote would be better than a sincere one. Previous research on voting systems' susceptibility to *ex post* manipulation thus has little to say about their susceptibility to *ex ante* strategic voting.

To see the difference uncertainty makes, consider a three-candidate plurality election in which the three candidates are expected to have roughly equal support. To evaluate manipulability using the typical *ex post* approach, we would first draw a large number of simulated elections from a distribution capturing our beliefs about likely outcomes; we would then count the proportion of cases in which an insincere vote could yield a better outcome for some voter. Given the assumption that the three candidates have roughly equal support, we would expect a roughly equal number of ties for first between each pair of candidates, each of which provides an opportunity for a voter whose first choice finishes third to obtain a better outcome by voting for their second choice; the more such ties, the more manipulable the system would appear to be. But when we consider the problem from the *ex ante* perspective, we arrive at a very different conclusion. If a tie for first is equally likely between each pair of candidates, a sincere vote is always optimal; fundamentally, this is because ties that reward sincerity are more numerous and have higher stakes on average than ties that reward insincerity.[4] The possibility of manipulable results means little for strategic voting unless, given reasonable uncertainty, a voter could discern that a result that would reward an insincere vote is substantially more likely than a result that would reward a sincere vote.

In this paper, we propose a general approach to assessing strategic voting incentives in the presence of uncertainty and use it to assess whether plurality or instant-runoff elections are more susceptible to strategic voting (the question on which Clegg, Cox, and Dummett all offered views above). Our measure of strategic voting incentives in a given electoral system is the answer to the question, "How much should a consequentialist voter (i.e. one who cares only about election outcomes) be willing to pay to vote strategically rather than sincerely?" The answer to this question depends on what we assume about voters' *preferences* and *beliefs*. For

[4]A sufficient condition for a sincere vote to be optimal in a three-candidate plurality election is that, given preference ordering *abc*, a *bc* tie and an *ac* tie are equally likely.

preferences, we use election surveys in the Comparative Study of Electoral Systems (CSES) in which voters are asked to rate each party on a 0-10 scale; this yields over 230,000 sets of preferences in 160 different elections.[5] Next, we model voters' beliefs about possible election outcomes as a probability distribution satisfying two criteria: first, the precision of the distribution is consistent with the empirical predictability of election outcomes; second, the location of the distribution is consistent the preference data (i.e. what other voters in the same election want) and one of a range of assumptions about the prevalence of strategic voting. This approach allows us to measure, for each voter and a given assumption about how strategic *other* voters are expected to be, the expected benefit of strategic voting compared to sincere voting and, ultimately, to compare this benefit across different voting systems.

We find that, consistent with some of the conjectures noted above, the incentive to vote strategically is considerably lower in IRV than in plurality elections. Decomposing the expected benefit of voting strategically, we observe that IRV is more resistant to strategic voting both because the probability of benefiting from a strategic vote is lower and because the magnitude of that benefit (when it exists) is smaller. We find that the expected benefit from strategic voting is lower in IRV regardless of how strategic other voters are expected to be, but the gap between plurality and IRV is larger the more strategic other voters are expected to be: as other voters become more strategic, the probability of benefiting from a strategic vote increases in plurality (because strategically deserting a *trailing* candidate becomes more attractive when other voters are expected to do so) but decreases in IRV (because strategically deserting a *leading* candidate becomes less attractive when other voters are expected to do so).

We emphasize that our focus in this paper is on the *incentive* to vote strategically given reasonable uncertainty about others' votes, not the empirical prevalence or practicability of strategic voting. In order to detect the strategic incentives we measure in this paper, a voter needs basic information about the likely prevalence of each ballot type, a good understanding of the voting system, and the ability to reason strategically. Many voters may lack one or more of these ingredients – especially in IRV, which previous researchers have noted is more complicated to manipulate than plurality and other voting systems.[6] In this sense, our paper sheds light

[5]In each survey we use preferences over the top three parties (in terms of national vote share) only.

[6]In a framework where voters are certain about each others' votes, Bartholdi, Tovey and Trick (1989) use

on strategic voting incentives as these incentives might be computed by a voting advice service or perceived by a highly sophisticated party strategist who has well-formed (though imprecise) beliefs about election outcomes and can work through the implications of alternative voting strategies in a given system. Future research can assess whether actual voters can perceive and act on these incentives, but this paper provides insight on whether these incentives exist, how strong they are, in what direction they point, and how they are likely to affect outcomes.

2 A framework for measuring strategic voting incentives

2.1 Non-technical overview

A sincere voter chooses the ballot that best reflects her sincere preferences; a strategic voter chooses the ballot that maximizes her expected utility given her preferences over the competing candidates and her beliefs about how others are likely to vote. Our objective is to measure the expected benefit of strategic voting in a given system, i.e. how much better off a typical voter is by being strategic rather than sincere, and to compare this across voting systems.

Consider first how to measure the benefit of strategic voting for a single voter. We need two inputs. First we need a measure of the voter's *preferences* in the form of a numerical utility from electing each candidate; we will use ratings voters assign to each party, but other measures are possible. Second we need a model of the voter's *beliefs* about possible election outcomes in the form of a probability distribution; we will use a Dirichlet distribution but again other approaches are possible. Given these two inputs, we can compute the expected utility for the voter from casting each possible ballot.[7] The benefit of strategic voting for this voter is the difference between the expected utility from casting a strategic vote and the expected utility from casting a sincere vote. Because the strategic vote by definition maximizes expected utility,

computational complexity theory to assess the difficulty of manipulating an outcome in various voting systems; Bartholdi and Orlin (1991) argues that manipulation in IRV is uniquely complex. Cox (1997) and Renwick (2011) also note that strategies in IRV are more logically complicated than strategies in plurality, which suggests that some voters may find it too difficult to even attempt.

[7] One way of computing the expected utility of a particular ballot is to shift the voter's beliefs by one vote in the direction of that ballot, compute the probability of each candidate winning given that shift, and compute the cross product of these probabilities and the utility measures. Another approach is to compute the probability of all *pivot events* (outcomes in which a single vote could determine the winner) given the beliefs and measure the expected utility of each ballot conditional on such a pivot event occurring, which is proportional to the expected utility of the ballot plus a constant.

this difference is 0 if the strategic vote is the same as the sincere vote and positive otherwise.

The key question in computing the expected benefit of strategic voting for the typical voter is what to assume about the inputs, i.e. the preferences and beliefs. Ideally, we might run an experiment in which we randomly assign a voting system across a very large number of polities, allow time for voters and candidates to adapt to the system, and measure the joint distribution of preferences and beliefs in the electorate under each voting system; we could then compute the expected benefit of strategic voting averaged over this joint distribution to characterize strategic voting incentives in each system. Unfortunately, not only is this experiment impossible, but the observational equivalent is impractical as well: IRV is not used widely enough for us to credibly infer a likely joint distribution of preferences and beliefs typical of that system.

Instead, we draw preference data for both IRV and plurality from a large set of electoral surveys (and thus our analysis measures the first-order effect of the voting system on strategic voting incentives while holding preferences fixed) and for each survey we generate a sequence of plausible beliefs by varying the extent to which voters are assumed to be strategic. More specifically, our preference data comes from a set of 160 electoral surveys from 59 countries in which voters are asked to rate the largest three parties in an election on a 0-10 scale;[8] we use these ratings as proxies for utility. We model beliefs about voting results using a Dirichlet distribution that is characterized by a precision parameter and an expected result (i.e. a share for each possible ballot). We allow the precision parameter to vary between the level of belief precision characteristic of voters (according to Fisher and Myatt, 2017) and of forecasters (according to Eggers and Vivyan, 2018). For each value of the precision parameter, we generate a sequence of expected results through a procedure that mirrors how a system might move toward equilibrium through a series of polls or elections: at the start voters are assumed to vote sincerely, but at each subsequent step a share of the voters strategically responds to the result of the previous step, such that each step can be viewed as an increase in the strategic-ness of the electorate. We then calculate the expected benefit of strategic voting at each point along this sequence (and thus with different assumptions about the prevalence of strategic voting) and compare these across voting systems.

[8] Voters are asked to rate more than three parties in most cases, but we use only the top three parties for simplicity and consistency.

In sum, we offer a precise definition of strategic voting incentives and a sensible approach to measuring these incentives. Our approach relies on preference data from electoral surveys and allows for a variety of assumptions about how precise voters' beliefs are and how strategically they behave.

In the rest of this section we formalize the framework.

2.2 Expected utility of a ballot

Candidates, ballots, and beliefs over voting results: We consider an election with K candidates (indexed by k) and B admissible ballots (indexed by b). Let $\mathbf{v}$ denote a B-length vector indicating the proportion of voters submitting each admissible ballot, with $\mathcal{V}$ denoting the set of all possible voting results. Let f represent beliefs about possible voting results, with $f(\mathbf{v})$ denoting the perceived probability of result $\mathbf{v}$.

Election outcomes and beliefs over election outcomes as a function of the marginal ballot: Let $w(k \mid b, \mathbf{v})$ be an indicator function that is 1 if, under the voting system being considered, the kth candidate wins when the marginal voter submits the bth ballot[9] and the result among other voters is $\mathbf{v}$. Then

$$p_{k,b} \equiv \sum_{\mathbf{v} \in \mathcal{V}} f(\mathbf{v}) w(k \mid b, \mathbf{v})$$

indicates the probability that the kth candidate is elected when the marginal voter submits the bth ballot. Let $\mathbf{P}$ denote a $K \times B$ matrix with representative element $p_{k,b}$. The bth column of $\mathbf{P}$ thus indicates the probability of each candidate being elected when the marginal voter submits the bth ballot.

Expected utility as a function of the marginal ballot: Let $u(k)$ indicate a voter's cardinal (i.e. von Neumann-Morgenstern) utility from electing the kth candidate, and let $\mathbf{u} =$

[9] Recall that b indexes the admissible ballots, so the bth ballot could be e.g. a vote for candidate a in a plurality election or a ballot that ranks the candidates abc in an IRV election.

$\{u(1), u(2), \ldots, u(K)\}$ indicate the K-length vector of these utilities. Then the B-length vector

$$\overline{\mathbf{u}} \equiv \mathbf{u}\mathbf{P} = \{\overline{u}(1), \overline{u}(2), \ldots, \overline{u}(B)\}$$

indicates the voter's expected utility from submitting each possible ballot. *Strategic voting* implies choosing the ballot that yields the highest expected utility, i.e. solving

$$\arg\max_b \overline{u}(b).$$

2.3 Measuring strategic voting incentives

Let b_s indicate the sincere ballot for a given voter, i.e. one that is consistent with $\mathbf{u}$,[10] and let b^* indicate the ballot that maximizes expected utility, i.e. the ballot that maximizes $\overline{u}(b)$. For a given combination of preferences $\mathbf{u}$ and beliefs $\mathbf{P}$ producing expected utility vector $\overline{\mathbf{u}}$, sincere voting yields $\overline{u}(b_s)$ while strategic voting yields $\overline{u}(b^*)$. Given a joint distribution of preferences and beliefs $g(\mathbf{u}, \mathbf{P})$, the *ex ante* expected benefit of strategic voting compared to sincere voting is

$$E\left[\overline{u}(b^*) - \overline{u}(b_s)\right] = \int\int \big(\overline{u}(b^*) - \overline{u}(b_s)\big) g(\mathbf{u}, \mathbf{P}) \,\mathrm{d}\mathbf{u}\,\mathrm{d}\mathbf{P}.$$

2.4 Specifying beliefs and preferences: baseline approach

To characterize typical strategic voting incentives in a given voting system, one needs to specify $g(\mathbf{u}, \mathbf{P})$, i.e. the joint distribution of beliefs $\mathbf{P}$ and preferences $\mathbf{u}$. Ideally this joint distribution reflects "typical" voting situations in a given system. Our approach derives both preferences and beliefs from election surveys in which voters are asked to rate the competing parties: we use survey respondents' party ratings as cardinal utilities, and we consider a sequence of assumptions about strategic voting to generate plausible expected election outcomes. Details follow.

Preferences: Our proxy for voter i's cardinal utility, $\mathbf{u}_i$, is drawn from the party ratings provided by election survey respondents. In the surveys we use, these ratings are responses to a

[10] The definition of a sincere vote depends on the voting system: in the plurality system a sincere vote is a ballot that indicate's the voter's most preferred candidate; in a ranking system like IRV a sincere vote is a ballot that lists the candidates in the voter's order of preference.

question like, "On a scale of 0 to 10 (where 0 means 'strongly dislike' and 10 means 'strongly like'), how do you feel about Party A?"[11]

Beliefs: The objective is to specify a model of beliefs over voting results that is reasonable both in terms of precision and location. Following Fisher and Myatt (2017) and Eggers and Vivyan (2018), we model beliefs as a Dirichlet distribution which, given B possible ballots, assigns a density to every point on a simplex with $B-1$ dimensions.[12] Specifically, the density at voting result $\mathbf{v} = \{v_1, v_2, \ldots, v_B\}$ given parameter vector $\alpha = \{\alpha_1, \alpha_2, \ldots, \alpha_B\}$ is

$$f(\mathbf{v}; \alpha) = \frac{\Gamma\left(\sum_{i=1}^{B} \alpha_i\right)}{\prod_{i=1}^{B} \Gamma(\alpha_i)} \prod_{i=1}^{B} v_i^{\alpha_i - 1}.$$

Following Fisher and Myatt (2017), we parameterize α such that $\alpha_i = s\overline{v}_i$, where s is the precision parameter and $\overline{v}_i$ is the expected share for ballot type i.

The analysis below uses a range of precision parameters s informed by recent empirical work on strategic voting. At the low end, we use $s = 10$, which Fisher and Myatt (2017) argue rationalizes strategic voting in recent plurality elections in England; that is, the likelihood of reported vote choice is maximized if voters' beliefs are modeled as a Dirichlet centered on the actual (or forecast) result with precision parameter $s = 10$. At the high end, we use $s = 85$, which Eggers and Vivyan (2018) argue characterizes constituency-level forecasts in recent plurality elections in the UK; that is, the likelihood of observed results is maximized if forecasters' beliefs are modeled as a Dirichlet centered on the published forecast with precision parameter $s = 85$.

To complete our model of beliefs we must choose $\overline{\mathbf{v}} \equiv \{v_1, v_2, \ldots, v_B\}$, the expected share

[11] When we estimate the benefit of strategic voting in one system and compare it to the benefit of strategic voting in another, our analysis is in terms of this like/dislike scale. For our comparisons to be meaningful, then, we rely on some comparability in the scale of the ratings across voters: it cannot be, for example, that the voters who would benefit from strategic voting in IRV tend to use a very narrow range of ratings while those who would benefit from strategic voting in plurality use a wide range, despite some fundamental similarity in the intensity of their preferences.

[12] Fisher and Myatt (2017) used a Dirichlet distribution to characterize aggregate uncertainty over vote shares in a three-candidate plurality election; Eggers and Vivyan (2018) extended that approach to K-candidate plurality elections. Note that the Dirichlet is a pdf and not a pmf, so the probability of any particular vote share result is 0; one can nevertheless estimate the probability of a particular result using Dirichlet beliefs by multiplying the density at a point by a normalizing constant representing the volume of continuous vote shares that are closest to the point, given the size of the electorate.

for each ballot type. Our objective is to pick an expected result $\overline{\mathbf{v}}$ that is realistic given the distribution of preferences and what we know about voter behavior. The problem is that existing research does not tell us what election result to expect given a particular distribution of preferences even in the simplest case of three-candidate plurality elections: there are multiple equilibria if a sufficient proportion of voters are strategic (e.g. Myerson and Weber, 1993) and empirical work varies widely in its estimates of how strategic voters are (e.g. Kawai and Watanabe, 2013; Fisher and Myatt, 2017). Our approach is to identify a sequence of expected voting results that captures how the election would turn out for a given distribution of preferences as we vary the assumed strategic-ness of the electorate from one extreme (fully sincere voting) to another (fully strategic voting). We then compute the expected benefit of strategic voting at each of these expected voting results, which allows us to assess strategic voting incentives under a range of assumptions about other voters' strategies.

In more detail, we map out a sequence of expected results $\overline{\mathbf{v}}_0, \overline{\mathbf{v}}_1, \ldots, \overline{\mathbf{v}}_m$ that is constructed as follows. The first expected result, $\overline{\mathbf{v}}_0$, is simply the proportion of survey respondents holding each sincere preference, i.e. the expected result if all voters vote sincerely. Let $\mathbf{v}^*(\overline{\mathbf{v}})$ indicate the voting result when voters best-respond given expected result $\overline{\mathbf{v}}$ (where implicitly $\mathbf{v}^*(\overline{\mathbf{v}})$ is also a function of preferences $\mathbf{U}$ and belief precision s). Subsequent expected results in the sequence are given by $\mathbf{v}_j = \lambda \mathbf{v}^*(\mathbf{v}_{j-1}) + (1-\lambda)\mathbf{v}_{j-1}$, i.e. they are the weighted average of the previous expected result and voters' best response to the previous expected result.

This sequence of expected results can be interpreted in terms of levels of rationality. By conventional usage, a level-0 voter is non-strategic; a level-1 voter votes strategically but believes others are level-0 voters; a level-2 voter votes strategically and thinks that other voters are either level-1 or level-0 voters; and so on. Thus $\overline{\mathbf{v}}_0$ can be thought of as the voting result among level-0 voters and $\mathbf{v}^*(\overline{\mathbf{v}}_0)$ can be thought of as the voting result among level-1 voters. Going one step further, $\mathbf{v}^*(\overline{\mathbf{v}}_1)$ can be thought of as the voting result among level-2 voters who expect a proportion λ of other voters to be level-1 and the remainder to be level-0 (because $\mathbf{v}_1 = \lambda \mathbf{v}^*(\overline{\mathbf{v}}_0) + (1-\lambda)\overline{\mathbf{v}}_0$). Continuing this logic, we can decompose the expected result at any point along the sequence into the weighted average of the voting results among voters with different levels of rationality. In particular, we can write $\overline{\mathbf{v}}_j$ as

$$\begin{aligned}\overline{\mathbf{v}}_j &\equiv \lambda \mathbf{v}^*(\mathbf{v}_{j-1}) + (1-\lambda)\mathbf{v}_{j-1} \\ &= \lambda \mathbf{v}^*(\overline{\mathbf{v}}_{j-1}) + (1-\lambda)\lambda \mathbf{v}^*(\overline{\mathbf{v}}_{j-2}) + (1-\lambda)^2 \lambda \mathbf{v}^*(\overline{\mathbf{v}}_{j-3}) + \ldots + (1-\lambda)^j \overline{\mathbf{v}}_0 \\ &= \lambda \sum_{k=1}^{j-1} (1-\lambda)^{j-k} \mathbf{v}^*(\overline{\mathbf{v}}_{k-1}) + (1-\lambda)^j \overline{\mathbf{v}}_0,\end{aligned}$$

where $\mathbf{v}^*(\overline{\mathbf{v}}_{k-1})$ is the voting result among level-k voters. Thus as we move along the sequence of expected results we assume higher and higher levels of rationality in the electorate: at step 0 all voters are level-0 (non-strategic) voters, while at step j only $(1-\lambda)^j$ are level-0, λ are level-j-1, and other voters have intermediate levels of rationality.

The sequence of expected results can also be interpreted as a *tatonnement*-like process to identify a strategic voting equilibrium. Let a strategic voting equilibrium refer to voting result $\overline{\mathbf{v}}$ such that $\mathbf{v}^*(\overline{\mathbf{v}}) = \overline{\mathbf{v}}$, i.e. when voters expect $\overline{\mathbf{v}}$ and cast optimal votes accordingly, the expected result is $\overline{\mathbf{v}}$. It follows that if our sequence of expected results arrives at an absorbing state (i.e. stays at the same value indefinitely), then it must be a strategic voting equilibrium. We can thus think of our procedure as a way of arriving at an equilibrium from a starting point of sincere voting. In the plurality case, for example, the procedure leads us from the sincere result to the closest Duvergerian equilibrium: if A and B receive the most sincere support, then the procedure will lead us to the equilibrium where A and B receive essentially all support (with any votes for other candidates coming from voters who are almost indifferent between A and B).[13] Leading us to an equilibrium is valuable both in the plurality case, where we know the set of equilibria but need a way of selecting one, and in IRV, where there is no characterization of equilibria for arbitrary distributions of preferences. To the extent that the path can be viewed as the process of learning about equilibrium, the beliefs at intermediate steps can also be viewed as relevant.

[13]The expected order of finish can also change as the sequence proceeds, e.g. if the expected order of finish with sincere voting is ABC but A and B's supporters tend to rank C last while C's supporters are nearly indifferent between A and B (as in the classic "divided-majority" scenario), in which case A and C may win essentially all votes in equilibrium.

2.5 Computing election probabilities as a function of the marginal ballot

As explained above, given preferences $\mathbf{u}$ we can identify the optimal vote and measure tactical voting incentives by calculating $\mathbf{P}$, which encodes the probability of each candidate being elected as a function of the marginal ballot. We could approximate $\mathbf{P}$ in a brute-force manner by drawing a large number of election results from the belief distribution and computing the proportion of times each candidate is elected when a single ballot of each type is added to the simulated result. The problem is that (given the low probability of a single ballot being decisive) it is computationally costly to get precise results, particularly as the number of possible ballots B becomes large (e.g. in ranking systems).

We can more easily obtain precise estimates if, rather than estimating the election probability matrix $\mathbf{P}$, we estimate only the part of $\mathbf{P}$ that varies with the marginal ballot cast. To see this, consider the case of a plurality election involving candidates $\{a, b, c\}$. Let π_a denote the probability that a wins by more than one vote (so that the marginal ballot is irrelevant, with π_b and π_c defined correspondingly) and let π_{ab} denote the probability that a and b finish tied for first (with π_{ac} and π_{bc} defined correspondingly). Assume that ties are decided by alphabetic precedence (so e.g. a beats b if they tie for first) and that any candidate is just as likely to win by one vote over another candidate as to tie for first with that candidate.[14] Then we can write

$$\mathbf{P} = \begin{bmatrix} \pi_a + 2(\pi_{ab} + \pi_{ac}) & \pi_a + \pi_{ac} & \pi_a + \pi_{ab} \\ \pi_b + \pi_{bc} & \pi_b + 2(\pi_{ab} + \pi_{bc}) & \pi_b + \pi_{ab} \\ \pi_c + \pi_{bc} & \pi_c + \pi_{ac} & \pi_c + 2(\pi_{ac} + \pi_{bc}) \end{bmatrix}$$

For example, $p_{a,a}$ (the top-left item) indicates that a wins given a marginal ballot of a if (i) a would win given any marginal ballot (π_a), (ii) a is tied with either b or c for first ($\pi_{ab} + \pi_{ac}$), or (iii) a trails b or c by one vote (again, $\pi_{ab} + \pi_{ac}$); similarly, $p_{a,b}$ (the top-center item) indicates that a wins given a marginal ballot of b if (i) a would win given any marginal ballot (π_a) or (ii)

[14] Both assumptions are common (e.g. Bouton, 2013; Fisher and Myatt, 2017) and simplify the exposition.

a is tied with c for first (π_{ac}). Next we can separate $\mathbf{P}$ into two parts

$$\mathbf{P} = \begin{bmatrix} 2(\pi_{ab} + \pi_{ac}) & \pi_{ac} & \pi_{ab} \\ \pi_{bc} & 2(\pi_{ab} + \pi_{bc}) & \pi_{ab} \\ \pi_{bc} & \pi_{ac} & 2(\pi_{ac} + \pi_{bc}) \end{bmatrix} + \begin{bmatrix} \pi_a & \pi_a & \pi_a \\ \pi_b & \pi_b & \pi_b \\ \pi_c & \pi_c & \pi_c \end{bmatrix} = \tilde{\mathbf{P}} + \mathbf{Q}.$$

Then the vector of expected utilities from each ballot can be expressed

$$\overline{\mathbf{u}} = \mathbf{u}\mathbf{P} = \mathbf{u}\tilde{\mathbf{P}} + \mathbf{u}\mathbf{Q} = \mathbf{u}\tilde{\mathbf{P}} + \text{constant},$$

where the last equality obtains because the columns of $\mathbf{Q}$ are identical by definition. The constant drops out when we compute the expected benefit of strategic voting (i.e. the difference between the optimal vote and the sincere vote).

One can use either simulations or analytical approaches to estimate $\tilde{\mathbf{P}}$. Following Fisher and Myatt (2017), we compute the elements of $\tilde{\mathbf{P}}$ by integrating the Dirichlet distribution along the relevant vote share loci. For example, in the plurality case we compute π_{ab} by integrating along the locus where $v_a = v_b > v_c$ and multiplying by $\frac{1}{N}$, where N is the size of the electorate (because a single ballot for a could push a ahead of b when $v_b - v_a \in \left[0, \frac{1}{N}\right]$). That is,

$$\pi_{ab} = \frac{1}{N} \int_{x=\frac{1}{3}}^{\frac{1}{2}} \mathrm{Dir}\big(x, x, 1-2x; s\overline{\mathbf{v}}\big) dx,$$

where $\mathrm{Dir}\big(x, x, 1-2x; s\overline{\mathbf{v}}\big)$ is the Dirichlet density with parameter vector $\gamma\overline{\mathbf{v}}$ evaluated where $v_a = x$, $v_b = x$, and $v_c = 1-2x$. The procedure is analogous for π_{ac} and π_{bc}. We use a similar procedure for computing $\tilde{\mathbf{P}}$ in IRV elections, although it is somewhat more difficult because there are more ballot types (six instead of three, for a three-candidate race) and more ways in which the marginal ballot can affect the outcome (twelve instead of three). In the Appendix we document our analytical/numerical approach to computing pivotal probabilities in IRV and compare it to a simulation-based approach; we show that the two methods agree closely but obtaining precise estimates given Dirichlet beliefs is much less computationally costly with our analytical/numerical approach.

3 Strategic voting in plurality and IRV

3.1 Qualitative characterization of incentives conditional on beliefs about others' strategies

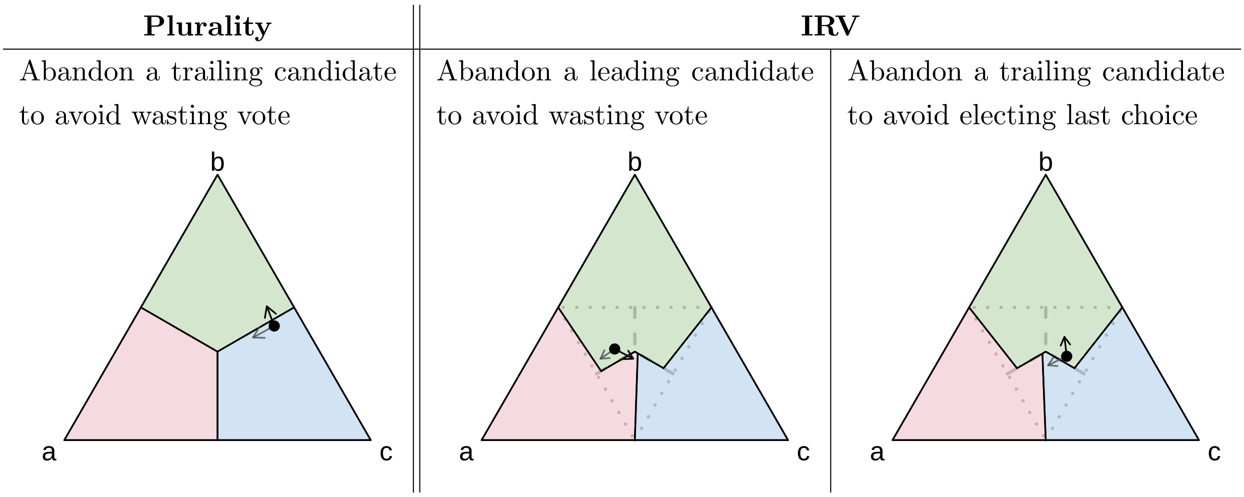
3.1.1 Plurality

Qualitative characterization Strategic voting incentives in plurality elections are simple and well understood. In a plurality election with candidates a, b and c, there are three *pivot events*, i.e. scenarios in which a single vote can determine who wins: a tie for first between a and b, a tie for first between a and c, and a tie for first between b and c. There are three possible votes (a, b, and c). For a voter whose sincere preference ordering of the candidates is abc, the optimal vote is a in the event of an ab or ac tie, but it is b in the event of a bc tie. Whether this voter should vote a or b thus depends on the relative probability of these events and the voter's preferences. With Dirichlet beliefs, a voter may optimally vote for her second choice either when her first choice is expect to finish third or when her first choice is expected to finish second behind her second choice (Fisher and Myatt, 2017). Thus the strategic voting incentive in plurality can be qualitatively described as "abandoning a trailing candidate to avoid wasting one's vote."

Visualization The strategic voting incentive in plurality is depicted on the ternary diagram at the bottom left of Figure 1. A ternary diagram depicts three shares that must sum to one (and thus that can be summarized by two numbers) in a way that maintains symmetry among the three alternatives.[15] The diagram is divided into three *win regions* indicating election results for which each candidate would win. The black dot indicates one possible result in which c narrowly defeats b, with a trailing far behind. Given that others' votes yield this result in expectation, how should a voter with preference order abc vote? A sincere vote for a would move the expected result toward a's vertex (following the gray arrow), while a vote for b would move the result toward b's vertex (following the black arrow). Given a reasonable distribution of uncertainty around the expected result, the latter move will have a bigger expected impact on the result and, to the extent that the voter prefers b over c, yield more in expected utility

[15] Katz and King (1999) provide an excellent introduction to the ternary diagram, which has been used to represent election results since at least Ibbetson (1965).

Figure 1: Types of strategic voting incentives in plurality and IRV elections



Note: There is one type of strategic voting incentive in plurality and two types in IRV. The ternary diagrams above depict a situation in which a voter with sincere preference *abc* might strategically vote *b* (left), *cab* (center), or *bac* (right). See Eggers (2019) for more on the use of ternary diagrams to represent IRV election results.

to the voter. Broadly, a wasted vote on the plurality ternary diagram is one that moves the expected result parallel to the nearest win-region boundary, while a strategic voter would cast a vote that moves the expected result perpendicular to the nearest win-region boundary.

3.1.2 IRV

Preliminaries The logic of strategic voting is more complex in an IRV election. To describe that logic, we will discuss a three-candidate IRV election as if it took place in two rounds: a first round in which the candidate with the fewest first-place votes is eliminated and a second round in which the winner is determined based on which of the remaining candidates is ranked higher on more ballots. A single vote can determine the winner at either stage: it can be *first-round pivotal* by determining which candidate is eliminated and thereby affecting which candidate wins, and it can be *second-round pivotal* by determining which of the non-eliminated candidates is elected.[16]

In total there are twelve pivot events in a three-candidate IRV election. There are three

[16]Technically it can also determine whether a candidate receives an outright majority of first-place rankings, but we can safely ignore these pivotal events because a candidate who is one vote away from winning an outright majority of first-place rankings could only lose the election if that candidate is not ranked second on any ballots that ranked the eliminated candidate first, which is unlikely in a large electorate.

scenarios in which a single ballot could determine the winner given that two candidates (say, a and b) are tied in the first round: c would lose to either a or b in the second round, c would lose to a but not b, or c would lose to b but not a. (If c would not lose to either a or b, then it is not a pivot event, as the winner does not depend on a single vote.) Thus for each of three pairs of candidates there are three first-round pivot events, for a total of nine first-round pivot events. Each pair of candidates can also be tied in the second round (e.g. c finishes last in the first round and, after c's votes are redistributed, a and b have the same level of support), for a total of three second-round pivot events and a grand total of twelve pivot events.

In principle we need to consider six possible ballots for each voter: *abc*, *acb*, *bac*, *bca*, *cab*, and *cba*.[17] In fact for any given voter we can focus on just three possible ballots.[18] To see why, note first that in any first-round pivot event the outcome depends only on one's top ranking, and as we will see shortly there are circumstances in which a voter benefits from giving the top ranking to her true first, second, or third choice. There is no reason to insincerely rank the other two candidates, however: these lower rankings are irrelevant in first-round pivot events and, when they matter for second-round pivot events, can only backfire if they do not reflect the voter's sincere preference. In short, whatever candidate one ranks first, one should rank the other two sincerely in case those candidates are tied in the second round. For each voter, then, the question is whether to give the top ranking to a, b, or c, with the lower rankings following the voter's sincere preference; for example, a voter with sincere preference *abc* should consider voting *abc*, *bac*, or *cab* but can ignore *acb*, *bca*, or *cba*.

Qualitative characterization In contrast to plurality, where there is one type of strategic

[17] If incomplete rankings are permitted, as they are in many Australian state legislative elections, then voters could also submit a ballot that ranks only two candidates (*ab*, *ac*, *ba*, *bc*, *ca*, and *cb*) or a ballot thats rank only one candidate (a, b, c), but for computing the optimal ballot we can ignore these incomplete rankings. If, when candidate e.g. a is eliminated, ballots listing a first are transferred to the candidate ranked second on those ballots, then a ballot ranking all but one candidate (e.g. *ab*) has the same effect on outcomes as a ballot that includes the omitted candidate at the end (e.g. *abc*), so we can consider those ballots as equivalent. If voters have strict preferences over candidates and all pivotal events have non-zero probability, then a ballot ranking only one candidate is strictly worse than a ballot ranking that candidate first followed by the voter's sincere ranking of the remaining candidates (because a second-round tie between those candidate is possible), so ranking a single candidate is never optimal. Fishburn and Brams (1984) showed that submitting a truncated ballot can be better than submitting a sincere ballot when there are four or more candidates, essentially because it avoids the no-show paradox, but even then another non-truncated ballot must be optimal.

[18] Dummett (1984, p. 224) makes this point.

voting incentive (which we described above as "abandoning a trailing candidate to avoid wasting one's vote"), in IRV there are two basic types of strategic voting incentive. The first type can be described as "abandoning a *leading* candidate to avoid wasting one's vote". This incentive arises when one's first choice, a, is expected to safely advance to the second round, but a tie for second between b and c is possible, such that by voting bac or cab one can determine whether b or c advances to the second round. Abandoning a leading candidate (here, a) is beneficial either when b and c would both defeat a in the second round (in which case voting bac or cab determines whether b or c is elected) or when a would beat one candidate (say, c) but not the other (in which case voting cab secures a's election).

The second type of strategic voting incentive in IRV can be described as "abandoning a trailing candidate to avoid electing one's least favorite candidate". This incentive arises when one's last choice (c) is expected to finish first in the first round, and one's second choice (b) can defeat c in the second round while one's first choice (a) cannot. (This might be the case if, for example, a, b, and c were arranged left to right on a single policy dimension, with b as the centrist and Condorcet winner.) If a and b were to tie for second, then a sincere abc vote would cause a to eliminate b, leading to the election of c, while a bac vote would cause b to advance and defeat c. By contrast with the strategic voting incentive in plurality, one abandons a trailing candidate here not because one's vote is *wasted*, but rather because it *backfires*; indeed, in this scenario one would be better off not voting than voting sincerely (which is why systems that produce this incentive are said to suffer from the "no-show paradox").

Visualization Following Eggers (2019) we extend the ternary diagram in Figure 1 to depict the two types of strategic voting incentives in IRV elections. The axes of the IRV ternary diagrams indicate the proportion of ballots ranking each candidate first. As in the plurality case, the triangle is divided into regions in which each candidate wins; unlike in the plurality case, the boundaries depend on the proportion of ballots ranking each candidate second conditional on which candidate is ranked first. In the IRV diagrams in Figure 1, the pattern of preferences is roughly single-peaked, with b as the centrist candidate.[19] Near the center of the diagram,

[19]See Eggers (2019) for more on the construction and interpretation of the ternary diagram for IRV.

the bottom boundary of b's win region follows an upside-down V shape. At results along the left branch of the upside-down V (highlighted in the center diagram), there is a tie for second between b and c in the first round, and b would win the election if b advanced while a would win the election if c advanced. For a voter with sincere preference *abc*, a result like this provokes the first type of IRV strategic voting incentive: a sincere *abc* vote would move the result toward a's vertex (as indicated by the gray arrow) but would not affect the winner, while a *cab* vote would move the result into a's win region (as indicated by the black arrow). At results along the right branch of the upside-down V (highlighted in the right diagram), there is a tie for second between a and b in the first round, and b would win the election if b advanced while c would win the election if a advanced. For a voter with sincere preference *abc*, a result like this provokes the second type of IRV strategic voting incentive: a sincere *abc* vote would move the result into c's win region, while a *bac* vote would move the result more safely into a's win region.

3.2 Dependence of strategic voting incentives on the level of strategic voting

Suppose a poll takes place in which voters are asked for their sincere ratings of competing candidates. What should a strategic voter believe about likely election results, given this polling information? Clearly the appropriate belief depends on the extent to which other voters are strategic and how these voters form their beliefs about likely election results. If other voters are completely non-strategic, the poll offers a good direct approximation of the likely election result; if other voters are themselves strategic, a strategic voter can only form a belief about likely results by first forming a belief about how others use the poll to form a belief. As discussed above, our analysis posits that voters have a range of levels of rationality; we examine strategic voting incentives for a sequence of beliefs as voters are perceived to become more strategic. For now, we seek only to develop intuition about how strategic voting incentives depend on beliefs about others' strategic orientation and why this differs in plurality and IRV elections; to do this, it is sufficient to compare the incentives of level-1 and level-2 voters (i.e. those who believe that other voters are non-strategic and those who believe that some other voters are level-1) and observe how this comparison differs between the plurality case and the IRV case.

To begin, it is useful to note that the strategic voting incentive (i.e. the average difference in expected utility between a strategic vote and a sincere vote) can be decomposed into two parts, which we will call *prevalence* and *magnitude*:

$$\underbrace{E\left[\overline{u}(b^*) - \overline{u}(b_s)\right]}_{\text{Exp. benefit of strategic voting}} = \underbrace{\Pr\left(\overline{u}(b^*) > \overline{u}(b_s)\right)}_{\substack{\text{Prob. tactical vote optimal} \\ \textbf{(prevalence)}}} \times \underbrace{E\left[\overline{u}(b^*) - \overline{u}(b_s) \mid \overline{u}(b^*) > \overline{u}(b_s)\right]}_{\substack{\text{Avg. benefit of an optimal tactical vote} \\ \textbf{(magnitude)}}}.$$

We focus in this section on the prevalence term; there is more ambiguity in how the magnitude of strategic voting incentives depends on beliefs about others' strategies. Our objective is to understand how the prevalence of strategic voting benefits depends on how strategic other voters are expected to be.

We begin with plurality. Recall that strategic voting in plurality elections involves abandoning candidates who are not expected to be in contention for first place. Level-1 voters expect the election result to resemble the poll in expectation. Given Dirichlet beliefs and three candidates, therefore, some level-1 voters abandon the candidate with the lowest sincere support;[20] the decision depends on belief precision and strength of relative preferences. If level-2 voters expect the election result to be a weighted average of the poll and the best responses of level-1 voters, and if strategic responses by level-1 voters make the candidate with the lowest sincere support even less competitive,[21] then the prevalence of strategic voting benefits (i.e. the proportion who optimally cast as insincere vote) should be higher among level-2 voters than level-1 voters. Because the incentive to desert an unpopular candidate tends to become more widespread when others are expected to desert that candidate, strategic voting in plurality elections is characterized by *strategic complementarity*: if others desert a trailing candidate, I am more likely to benefit from deserting that candidate.

In IRV the situation is different. Suppose that (as above) a is the left-wing candidate, b is the centrist candidate, and c is the right-wing candidate, and suppose that c has the largest share of sincere first-preferences. Given level-1 beliefs, a voter with cba preference order has

[20] Others may abandon the candidate with the second-lowest sincere support.

[21] This is not guaranteed because of desertions from the candidate with the second-lowest sincere support. For example, strategic votes by level-1 voters could reverse the expected order of finish of the sincere second- and third-place candidates, such that the candidate with the lowest sincere support is seen as more competitive by level-2 voters.

an incentive to submit a acb ballot to help the weaker opponent, a, advance to the second round. (a is expected to be the weaker opponent because a is listed second on only some of b's ballots while b is listed second on all of a's ballots.) For the same reason, a voter with abc preference order has an incentive to submit a bac ballot to avoid the situation where a eliminates b, electing c. (These are examples of the two strategic voting incentives in IRV.) Given level-2 beliefs, however, these incentives will be more limited. The incentive for cba types to vote acb is reduced in part because, as more voters desert the leading candidate to avoid wasting their vote, that candidate's lead becomes less secure. More subtly, both incentives are reduced because the strategic votes of type-1 voters, prompted by a's weakness relative to b as an opponent for c, tend to make candidate a *stronger* relative to b as an opponent to c, thus producing negative feedback. To see this, note that both $cba \rightarrow acb$ votes and $abc \rightarrow bac$ votes tend to increase the proportion of votes for a that list c second, and $abc \rightarrow bac$ votes tend to increase the proportion of votes for b that list a second. This diminishes the discrepancy that motivated these votes in the first place: as a result of these strategic moves, and conditional on a tie for second between a and b in the second round, a is a stronger potential opponent for c (because more of b's ballots list a second) and b is a weaker potential opponent for c (because more of a's ballots list c second). The general pattern in IRV is that, when voters strategically respond to a discrepancy between the second preferences of two candidates (here, b's second preferences being more favorable to c than a's are), this discrepancy tends to disappear. Because both incentives in IRV tend to diminish when others are expected to act on them, strategic voting in IRV elections is characterized by *strategic substitutability*: if others desert a leading candidate to avoid wasting a vote or desert a trailing candidate to avoid electing their least favorite candidate, I am less likely to follow suit.[22]

[22]There are exceptions. For example, deserting a leading candidate who cannot defeat either candidate in the second round can exhibit strategic complementarity: the less support the candidate gets, the more important it is to determine which other candidate advances. This incentive, too, must eventually attenuate, however, as the candidate becomes less and less likely to advance.

4 Data

To assess the prevalence and distribution of strategic incentives under Plurality and IRV empirically, we rely on data from the Comparative Study of Electoral Systems (CSES) for a realistic set of preferences and beliefs. The dataset covers 160 surveys from xx different countries, administered shortly before or after an election.[23] We focus on the three largest parties (evaluated how?) and label them A, B, C in descending size, respectively. From each survey in the data set, we take respondents' party like/dislike scores for these parties to approximate voters' ordinal utilities and, by extension, their preferences over the parties.

Let $\tilde{\mathbf{v}}$ be the vector of ballot proportions if everyone in the CSES survey voted sincerely according to their preferences. If voter i believes that everyone else is voting sincerely (in other words, everyone else is Level-0 strategic), then we model the i's belief about the next election as $\text{Dir}(s \times \tilde{\mathbf{v}})$, where s is a parameter capturing the precision of i's beliefs. Given this set-up of beliefs and preferences, we calculate the strategic incentives under either electoral system, and iterate this procedure as laid out in Section 3.

The remainder of this section describes brief summary statistics of the CSES data.

4.1 Summary statistics

The mean number of respondents in the CSES surveys is 1384 (with a standard deviation of 539). The 160 different surveys come from xx different countries, between 1996 and 2016. Figure **??** maps the number of surveys in each country. (Do we need to say any more? Perhaps something about mean / sd of preference intensity and τ?)

4.2 Weighting

In some CSES cases, respondents are assigned a non-uniform weight. When computing case-level statistics (e.g., prevalence of strategic incentives in case j), we weigh each observation by its original weight. When aggregating up even further, and presenting aggregate statistics (averaged over all cases), we also assign case-level weights to adjust for countries' voting age

[23] Two additional cases in the survey, Belarus (20xx) and Lithuania (20xx), are dropped because no respondent specified full preferences over more than two parties.

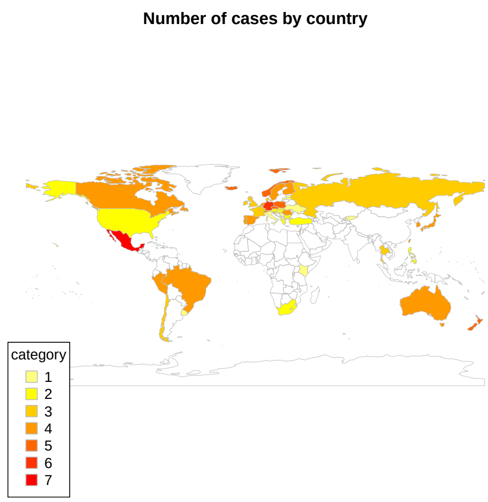


Figure 2: Cases in CSES data, by country

population and the overrepresenation of some countries.[24]

These case-level weights are constructed as follows:

$$w_j \equiv \tag{1}$$

4.3 Distribution of preferences

How different are the CSES cases from one another? Aside from the intensity of preferences, we can describe each case with the vector $\tilde{\mathbf{v}}$, where the three-item vector $(v_1 + v_2, v_3 + v_4, v_5 + v_6)$ describes the distribution of first preferences, and the three-item vector $(m_{AB} = \frac{v_1}{v_1+v_2}, m_{BA} = \frac{v_3}{v_3+v_4}, m_{CB} = \frac{v_6}{v_5+v_6})$ describes the distribution of second preferences.

To link these two distributions together and classify cases more completely, we offer the following approach. Without loss of generality, let the candidate (party) X whose first-preference voters have the most equally split second preferences, and the other two parties Y and Z. If both $m_{YZ}, m_{ZY} > 0.6$, then classify this case as *single-peaked* and denote it $X+$.[25] Conversely,

[24]Recall that our initial objective is to compare the *overall* distribution of strategic voting incentives under Plurality and IRV. Without weights, we would run the risk of having our findings distorted by a small country-outlier that counts for as much as a large state (e.g., Denmark and the United States); alternatively, we also do not want a result that is particular to one country to be over-represented purely because there are multiple surveys from that country.

[25]X is the attractor: both remaining parties have a majority of their second preferences tilted towards X.

Table 1: Distribution of preference profiles in CSES data

	A	**B**	**C**
Single-peaked (+)	18	23	9
Divided majority (-)	28	20	20
Neutral ()	5	7	3
Other ()		27	

if both $m_{YZ}, mZY < 0.4$, then classify this case as *divided majority* and denote it $X-$.[26] If $m_{YZ}, m_{ZY} \in [0.4, 0.6]$, then classify this case as *neutral* and denote it $N(X)$. If neither of these conditions hold (because of unusual second preferences), classify it as *other* and denote it O. This completes a mutually exclusive and exhaustive set of classes determined by $\tilde{\mathbf{v}}$.

Table 1 summarises the distribution of preference classes across the CSES cases. A plurality of cases belong to the divided majority classes; however, there is also a large number of single-peaked cases, whereas neutral and others tend to be rarer. (Figure 3 plots the distribution of first preferences conditional on the classes.)

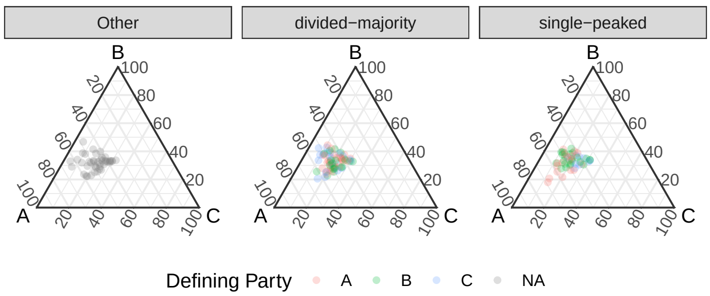


Figure 3: Distribution of first preferences in CSES cases, by class

5 Results

We now proceed to present and discuss our results.

[26]Here, X is the repeller: both remaining parties have a majority of their second preferences tilted towards each other and away from X.

5.1 Convergence

Under both IRV and Plurality, the distribution of ballot shares quickly converges towards a fixed point in the vast majority of CSES cases. The average Euclidean distance going from the 59th to the 60th iteration is below 0.0014 for Plurality, and below 0.006 for IRV.[27] Put differently, we can obtain a perfectly strategic voting equilibrium, where all voters anticipate others' vote choices, and react accordingly, within about 60 iterations from the sincere voting profile.

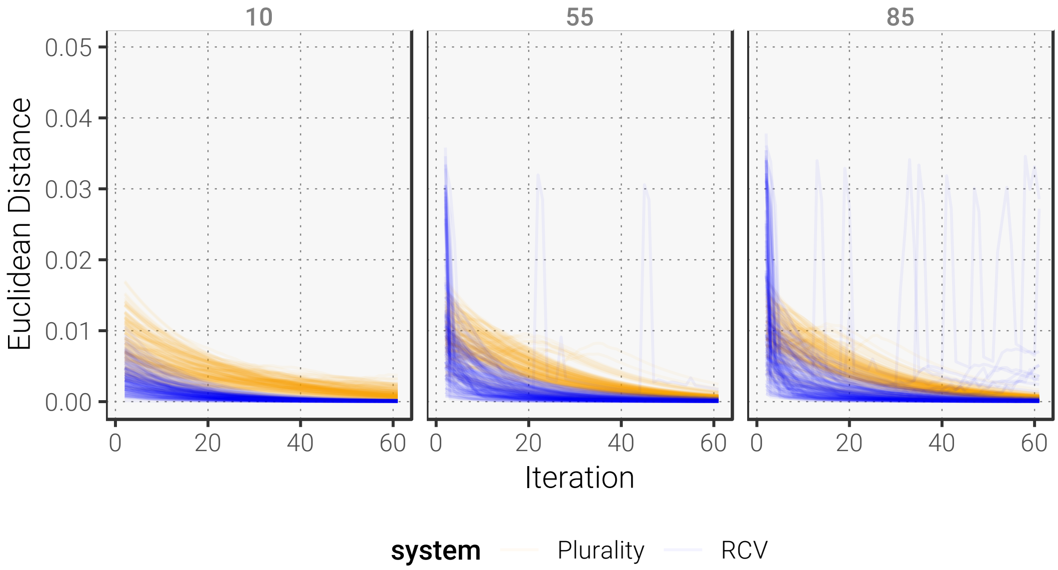


Figure 4: Euclidean distance between ballot share vectors from one iteration to another.

Figure 4 plots the Euclidean distance between the ballot shares for every case and iteration under both Plurality and IRV. In expectation, convergence towards the fixed point occurs faster under IRV than it does under Plurality. As we discussed earlier, strategic incentives under Plurality are characterised by complementarity; this means that with every additional iteration, the incentive for supporters of the third party increases, until all of them have deserted the trailing candidate and the ballot shares are in a Duvergian (two-party) equilibrium.[28] In contrast, the substitutability of strategic voting incentives under RCV allows them to reach a fixed point much sooner. Note however, that, for more precise beliefs ($s \in 55, 85$), the shift away from the sincere ballot profile in the first few iterations is much bigger than under Plurality;

[27] These averages are unweighted – need to recompile in the future.

[28] We could visualise this by plotting the share of third-party votes when $k = 60$.

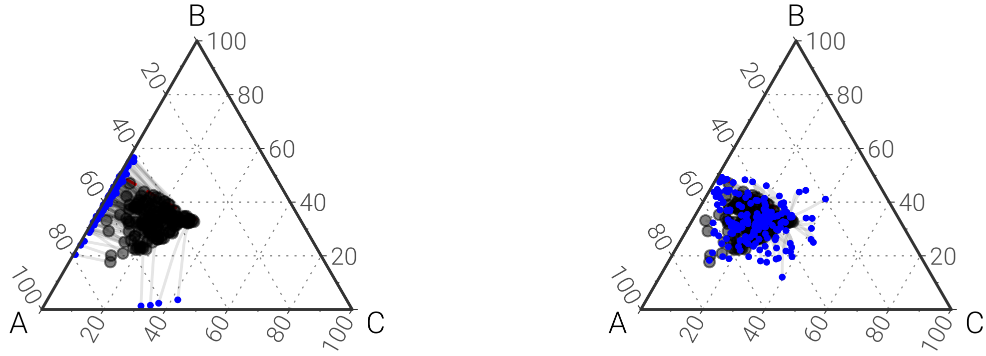


Figure 5: Evolution of ballot share vectors for all CSES cases over iterations, for both Plurality (left) and IRV (right), when $s = 85$. Grey dots indicate the initial ballot share vector before the first iteration; blue dots the ballot share vector after the 60th iteration.

quicker convergence does not necessarily mean that the fixed point is closer to the original ballot share vector.[29]

In sum, when applying our iterative strategic voting procedure to all CSES cases, the ballot shares converge more quickly to a fixed point under IRV than under Plurality. Under IRV, these fixed points can occur anywhere in the ballot share space, whereas under Plurality, voters ultimately settle on a two-party Duvergian equilibrium. This is also illustrated by Figure 5, which maps the ballot share vectors before the first and the 60th iteration for $s = 85$.

(Figure about distance from sincere profile? – shows nicely that Plurality fixed points are further away from initial ballot shares.)

5.2 Strategic Incentives

In this section, we present our main results. We focus on the prevalence, magnitude and expected benefit of strategic voting under either electoral system. Overall, strategic voting incentives are more prevalent, have a higher magnitude and higher expected benefit under Plurality than under IRV.

[29]This foreshadows a later result: with sufficiently high precision, the prevalence of strategic voting incentives under IRV will be higher in the first few incentives.

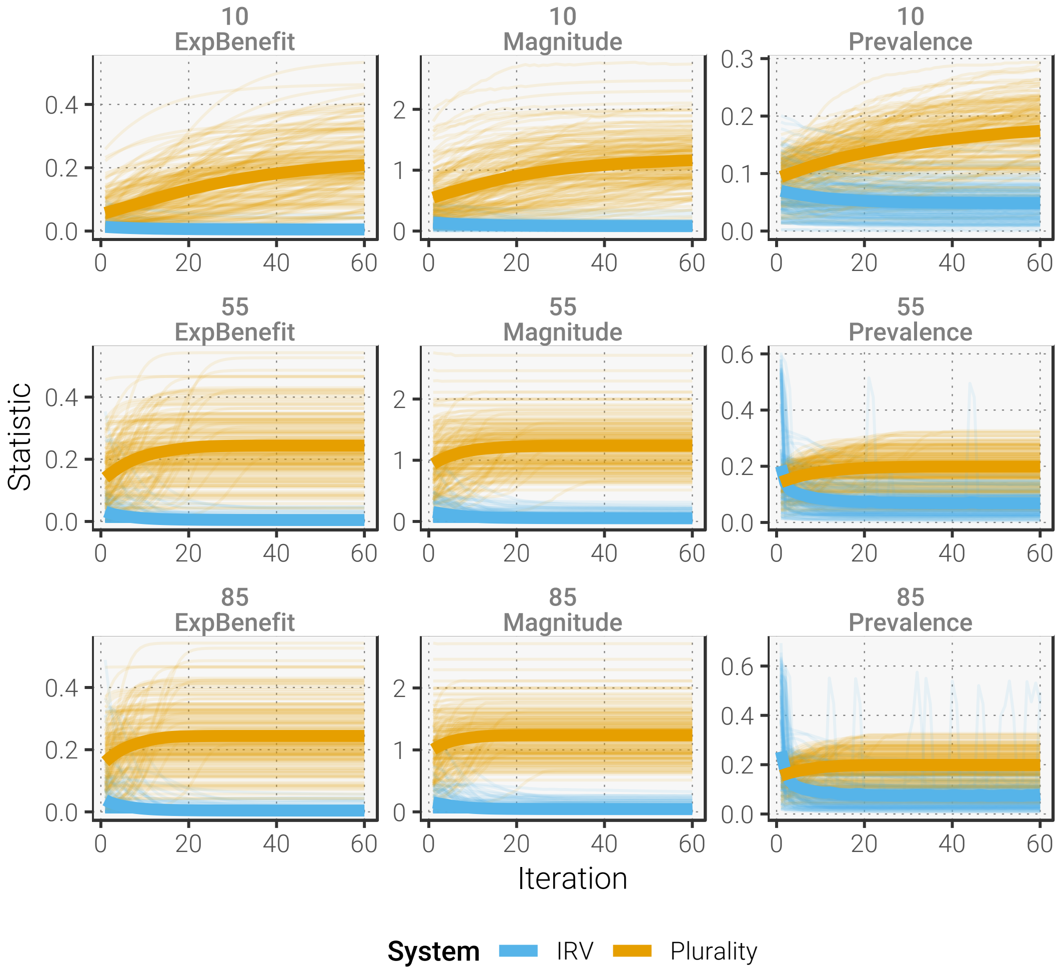


Figure 6: Main statistics

Figure 6 shows the quantities of interest for each case, as well as the weighted average.

6 Conclusion

References

Bartholdi, John J, Craig A Tovey and Michael A Trick. 1989. "The computational difficulty of manipulating an election." *Social Choice and Welfare* 6(3):227–241.

Bartholdi, John J and James B Orlin. 1991. "Single transferable vote resists strategic voting." *Social Choice and Welfare* 8(4):341–354.

Bouton, Laurent. 2013. "A theory of strategic voting in runoff elections." *American Economic Review* 103(4):1248–88.

Chamberlin, John R. 1985. "An investigation into the relative manipulability of four voting systems." *Behavioral Science* 30(4):195–203.

Cox, Gary W. 1997. *Making votes count: strategic coordination in the world's electoral systems.* Cambridge University Press.

Dummett, Michael. 1984. *Voting procedures.* Oxford University Press.

Eggers, A.C. 2019. "A diagram for analyzing elections that use ranked preferences.".

Eggers, Andrew C and Nick Vivyan. 2018. "Who votes more strategically?".

Favardin, Pierre and Dominique Lepelley. 2006. "Some further results on the manipulability of social choice rules." *Social Choice and Welfare* 26(3):485–509.

Fishburn, Peter C and Steven J Brams. 1984. "Manipulability of voting by sincere truncation of preferences." *Public Choice* 44(3):397–410.

Fisher, Stephen D and D Myatt. 2017. "Strategic voting in plurality rule elections." *Mimeo* .

Frigyik, Bela A, Amol Kapila and Maya R Gupta. 2010. Introduction to the Dirichlet distribution and related processes. UWEE Technical Report UWEETR-2010-0006 Department of Electrical Engineering, University of Washington.

Ibbetson, D. 1965. "Comment on Berrington, 'The general election of 1964'." *Journal of the Royal Statistical Society. Series A (General)* 128(1):54–55.

Katz, Jonathan N and Gary King. 1999. "A statistical model for multiparty electoral data." *American Political Science Review* 93(1):15–32.

Kawai, Kei and Yasutora Watanabe. 2013. "Inferring strategic voting." *American Economic Review* 103(2):624–62.

Miller, Nicholas R. 2017. "Closeness matters: monotonicity failure in IRV elections with three candidates." *Public Choice* 173(1-2):91–108.

Myerson, Roger B and Robert J Weber. 1993. "A theory of voting equilibria." *American Political Science Review* 87(1):102–114.

Nitzan, Shmuel. 1985. "The vulnerability of point-voting schemes to preference variation and strategic manipulation." *Public choice* 47(2):349–370.

Ornstein, Joseph T and Robert Z Norman. 2014. "Frequency of monotonicity failure under Instant Runoff Voting: estimates based on a spatial model of elections." *Public Choice* 161(1-2):1–9.

Plassmann, Florenz and T Nicolaus Tideman. 2014. "How frequently do different voting rules encounter voting paradoxes in three-candidate elections?" *Social Choice and Welfare* 42(1):31–75.

Renwick, Alan. 2011. The alternative vote: a briefing paper. Technical report.

Saari, Donald G. 1990. "Susceptibility to manipulation." *Public Choice* 64(1):21–41.

A The probability of pivotal events in IRV elections

Given beliefs about the likelihood of possible election outcomes, the probability of pivotal events (i.e. results such that the outcome could be affected by a single ballot) can be calculated either by a numerical/analytical method involving integration of the belief distribution or by a simulation method. This section presents both approaches (beginning with the numerical/analytical method) and shows that they produce the same result, though the simulation method is much more computationally intensive.

Ballots: In an IRV election with three candidates $\{a, b, c\}$ in which all ballots must rank all candidates,[30] there are six admissible ballots. Let v_{ab} denote the share of ballots ranking candidate a first, b second, and (implicitly) c third (with v_{ac}, v_{ba}, etc. defined equivalently). Then the vector of ballot shares is

$$\mathbf{v} = \{v_{ab}, v_{ac}, v_{ba}, v_{bc}, v_{ca}, v_{cb}\}.$$

Let $v_a \equiv v_{ab} + v_{ac}$ denote the share of ballots ranking candidate a first (with v_b and v_c defined equivalently); we will refer to v_a, v_b, and v_c as *first-preference shares*. Finally, let $\overline{v}_{ab}$ denote the expected share of ballots ranking a first and b second, with $\overline{\mathbf{v}}$ indicating the vector of expected ballot shares (and $\overline{v}_{ac}$, $\overline{v}_{ba}$, etc. defined equivalently) and let $\overline{v}_a \equiv \overline{v}_{ab} + \overline{v}_{ac}$ (with $\overline{v}_b$ and $\overline{v}_c$ defined equivalently).

Pivotal events: There are two broad classes of pivotal results in a three-candidate IRV election. In a *first-round pivotal event*, two candidates tie for second place in first-preference shares, and the identify of the winner depends on which candidate is eliminated. Let *ab.ab* denote the first-round pivotal event in which a and b tie for second place in first-preference shares and a wins the election if a advances while b wins the election if b advances; similarly, let *ab.ac* denote the first-round pivotal event in which a and b tie for second place and a wins the election if a advances while c wins the election if b advances; let *ab.cb* denote the first-round pivotal event in which a and b tie for second place and c wins the election if a advances while b wins the election if b advances. Let *ac.ac*, *ac.ab*, *ac.bc*, *bc.bc*, *bc.ba*, and *bc.ac* be defined similarly. In a *second-round pivotal event*, two candidates tie after the other candidate is eliminated; let *ab* denote the second-round pivotal event involving a and b, with *ac* and *bc* defined similarly. We will denote the probability of pivotal event e.g. *ab* by π_{ab}.

Beliefs: We assume that election outcomes are believed to follow a Dirichlet distribution centered on $\overline{\mathbf{v}}$ with precision captured by γ, i.e.

$$\mathbf{v} = \{v_{ab}, v_{ac}, v_{ba}, v_{bc}, v_{ca}, v_{cb}\} \sim \text{Dir}(\alpha_{ab}, \alpha_{ac}, \alpha_{ba}, \alpha_{bc}, \alpha_{ca}, \alpha_{cb}) \tag{2}$$

where e.g. $\alpha_{ab} = \gamma \overline{v}_{ab}$. We will make use of three well-known (Frigyik, Kapila and Gupta, 2010) properties of the Dirichlet distribution:

Aggregation property: $(v_1, v_2, \ldots, v_i + v_j, \ldots v_B) \sim \text{Dir}\big(\alpha_1, \alpha_2, \ldots, \alpha_i + \alpha_j, \ldots \alpha_B\big)$. (If two of the vote shares are added together to create a new, shorter vector of vote shares, the new vector of vote shares also follows a Dirichlet distribution, where the parameters corresponding to the summed-up vote shares are also summed up.)

Marginal distribution: $v_i \sim \text{Beta}(\alpha_i, \sum_{-i} \alpha)$. (Unconditionally, any particular vote share follows a Beta distribution. This follows from the aggregation property and the observation that a

[30] The analysis can be extended to allow truncated ballots.

Dirichlet distribution with two parameters is a Beta distribution.)

Conditional distribution: $(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_B \mid v_i) \sim (1-v_i)\text{Dir}(\alpha_1, \ldots, \alpha_{i-1}, \alpha_{i+1}, \ldots, \alpha_B)$. (Conditional on i receiving share v_i, the remaining shares follow a rescaled Dirichlet distribution in which α_i is removed from the parameter vector.)

We will use $f(\mathbf{v}; \gamma\overline{\mathbf{v}})$ to indicate the Dirichlet density with parameters $\gamma\overline{\mathbf{v}}$ evaluated at $\mathbf{v}$. Because the Beta density can be seen as a special case of the Dirichlet density, we will use $f(\cdot)$ for both.

Probability of second-round pivotal events: If we say that two candidates tie when their vote share differs by less than half a vote,[31] then the probability of second-round pivotal event ab can be written

$$\Pr\left(v_c < v_a < \frac{1}{2} \cap v_c < v_b < \frac{1}{2} \cap v_a + v_{ca} - \frac{1}{2} \in \left(-\frac{1}{2N}, \frac{1}{2N}\right)\right).$$

This can be factorized as

$$\Pr\left(v_a + v_{ca} - \frac{1}{2} \in \left[-\frac{1}{2N}, \frac{1}{2N}\right)\right) \times \Pr\left(v_c < v_a \cap v_c < v_b \,\middle|\, v_a + v_{ca} - \frac{1}{2} \in \left(-\frac{1}{2N}, \frac{1}{2N}\right)\right). \tag{3}$$

Using the aggregation property, the first term in expression 3 is

$$\int_{s=\frac{-1}{2N}}^{\frac{1}{2N}} \int_0^{\frac{1}{2}} f\left(y - s/2, \frac{1}{2} - y - s/2, \frac{1}{2} + s; \gamma\overline{v}_a, \gamma\overline{v}_{ca}, \gamma(\overline{v}_b + \overline{v}_{cb})\right) dy\, ds$$

which is approximately

$$\frac{1}{N} \int_0^{\frac{1}{2}} f\left(y, \frac{1}{2} - y, \frac{1}{2}; \gamma\overline{v}_a, \gamma\overline{v}_{ca}, \gamma(\overline{v}_b + \overline{v}_{cb})\right) dy.$$

(The approximation is exact if the density is flat in the immediate neighborhood of second-round ties between a and b.) We now turn to the second term in expression 3. Given that $v_a = y$, $v_{ca} = \frac{1}{2} - y$, and $v_b + v_{cb} = \frac{1}{2}$, we note that $v_c < v_a$ implies $v_{cb} < 2y - \frac{1}{2}$ and $v_c < v_b$ implies $v_{cb} < \frac{y}{2}$; comparing the two conditions, note that the former binds when $y < \frac{1}{3}$ and the latter binds otherwise. Next, using all three properties of the Dirichlet notes above and given that $v_a + v_{ca} = \frac{1}{2}$,

$$(v_{cb} \mid v_a + v_{ca}) \sim \frac{1}{2}\text{Beta}\left(\gamma\overline{v}_{cb}, \gamma\overline{v}_b\right), \tag{4}$$

i.e. given that half the ballots list a first or list c first and a second, the proportion listing c first and b second (instead of b first) lies between 0 and 1/2; if we multiply the proportion by two, the result is distributed according to a Beta distribution with parameters $\gamma\overline{v}_{cb}$ and $\gamma\overline{v}_b$. Thus to find the probability that $v_{cb} < 2y - \frac{1}{2}$ (the binding constraint in the second term from expression 3 when $y < 1/3$), we integrate this distribution from 0 to $2y - \frac{1}{2}$; to find the probability that $v_{cb} < \frac{y}{2}$ (the binding constraint in the second term from expression 3 when $y > 1/3$), we integrate this distribution from 0 to $\frac{y}{2}$. Finally note that y (i.e. v_a) cannot be below 1/4; otherwise either a finishes last in first-preference votes or b receives more than half

[31] This is equivalent to saying that we calculate the probability of specific results by rounding continuous vote shares to the closest multiples of $1/N$.

of first-preference votes. Combining all of this, we have

$$\begin{aligned}N\pi_{ab} \approx & \int_{\frac{1}{4}}^{\frac{1}{3}} f\left(y, \frac{1}{2}-y, \frac{1}{2}; \gamma\overline{v}_a, \gamma\overline{v}_{ca}, \gamma(\overline{v}_b+\overline{v}_{cb}\right) \int_0^{2y-\frac{1}{2}} f\left(2z, 1-2z; \gamma\overline{v}_{cb}, \gamma\overline{v}_b\right) dz\, dy + \\ & \int_{\frac{1}{3}}^{\frac{1}{2}} f\left(y, \frac{1}{2}-y, \frac{1}{2}; \gamma\overline{v}_a, \gamma\overline{v}_{ca}, \gamma(\overline{v}_b+\overline{v}_{cb}\right) \int_0^{\frac{y}{2}} f\left(2z, 1-2z; \gamma\overline{v}_{cb}, \gamma\overline{v}_b\right) dz\, dy. \end{aligned} \quad (5)$$

Note that the second and fourth densities are evaluated at $(v_{cb} = 2z, v_b = 1-2z)$ rather than $(v_{cb} = z, v_b = \frac{1}{2} - z)$ because of the $\frac{1}{2}$ in expression 4.

The analysis extends straightforwardly to the two other second-round pivotal events by exchanging candidate labels.

Probability of first-round pivotal events: First-round pivotal event $ab.ab$ takes place when a ties b for second place in first-preference votes and either candidate would win the election if the other were eliminated. Generally, the probability of $ab.ab$ is

$$\Pr\Big(v_b - v_a \in \Big(-\frac{1}{2N}, \frac{1}{2N}\Big) \cap v_b < v_c \cap v_a < v_c < \frac{1}{2} \cap v_a + v_{ba} > v_c + v_{bc} \cap v_b + v_{ab} > v_c + v_{ac}\Big), \quad (6)$$

which can be factorized as

$$\begin{aligned} &\Pr\Big(v_b - v_a \in \Big(-\frac{1}{2n}, \frac{1}{2n}\Big) \cap v_b < v_c \cap v_a < v_c < \frac{1}{2}\Big) \times \\ &\Pr\Big(v_a + v_{ba} > v_c + v_{bc} \cap v_b + v_{ab} > v_c + v_{ac} \Big| v_b - v_a \in \Big(-\frac{1}{2N}, \frac{1}{2N}\Big) \cap v_b < v_c \cap v_a < v_c < \frac{1}{2}\Big). \end{aligned}$$

Using the same approximation as above, the first line is approximately

$$\frac{1}{N} \int_{\frac{1}{4}}^{\frac{1}{3}} f\Big(z, z, 1-2z; \gamma\overline{v}_a, \gamma\overline{v}_b, \gamma\overline{v}_c\Big) dz.$$

Letting $v_a = v_b = z \in \left(\frac{1}{4}, \frac{1}{3}\right)$, the second term becomes

$$\Pr\Big(v_{bc} < 2z - \frac{1}{2} \cap v_{ac} < 2z - \frac{1}{2} \Big| v_a = v_b = z\Big). \quad (7)$$

and again combining all three properties we have

$$\begin{aligned} (v_{bc}|v_a + v_c) &\sim z\mathrm{Beta}\big(\gamma\overline{v}_{bc}, \gamma\overline{v}_{ba}\big) \\ (v_{ac}|v_b + v_c) &\sim z\mathrm{Beta}\big(\gamma\overline{v}_{ac}, \gamma\overline{v}_{ab}\big). \end{aligned}$$

Putting together the above, we have

$$\begin{aligned} N\pi_{ab.ab} \approx & \int_{\frac{1}{4}}^{\frac{1}{3}} f\Big(z, z, 1-2z; \gamma\overline{v}_a, \gamma\overline{v}_b, \gamma\overline{v}_c\Big) \times \\ & \int_0^{2z-\frac{1}{2}} f\left(\frac{x}{z}, \frac{z-x}{z}; \gamma\overline{v}_{bc}, \gamma\overline{v}_{ba}\right) dx \times \int_0^{2z-\frac{1}{2}} f\left(\frac{x}{z}, \frac{z-x}{z}; \gamma\overline{v}_{ac}, \gamma\overline{v}_{ab}\right) dx\, dz. \end{aligned} \quad (8)$$

To get the probability of pivotal event $ab.ac$ we reverse the last inequality in expression 6 (changing $v_b + v_{ab} > v_c + v_{ac}$ to $v_b + v_{ab} < v_c + v_{ac}$), which means changing the last term in expression 8 from $\int_0^{2z-\frac{1}{2}} f\left(\frac{x}{z}, \frac{z-x}{z}; \gamma\overline{v}_{ac}, \gamma\overline{v}_{ab}\right) dx$ to $1 - \int_0^{2z-\frac{1}{2}} f\left(\frac{x}{z}, \frac{z-x}{z}; \gamma\overline{v}_{ac}, \gamma\overline{v}_{ab}\right) dx$. The analysis extends straightforwardly to all other first-round pivotal events by similarly reversing inequalities and/or exchanging candidate labels.

Numerical estimation: We compute the probabilities above using numerical integration.

Simulation-based estimation: These probabilities can also be estimated directly via simulation by drawing M times from the belief distribution and counting the number of pivotal events. To make our estimates more computationally efficient, we count what we call *expanded pivotal events*, e.g. a and b receiving vote shares within δ rather than within $\frac{1}{2N}$ for some fixed N, and we divide by 2δ to yield an estimate of N times the probability of the pivotal event. The choice of δ reflects a bias-variance tradeoff: to the extent that the density is curved in the vicinity of the pivotal event, higher δ introduces bias into our estimates of pivotal probabilities, but the variance is roughly inversely proportional to δ.[32] The optimal choice of δ will depend on the cost of additional simulations and the shape of the density near pivotal events.

Checking consistency of numerical and simulation-based estimates: To check the validity of the numerical approach and compare the computational burden of the two approaches, we computed pivotal probabilities for 100 scenarios using the two approaches while varying the number of simulation draws. If our numerical approach is correct, the simulation results should converge on our numerical solutions as the number of simulations (and the computational burden of the simulation approach) increases. Below we show that this is the case.

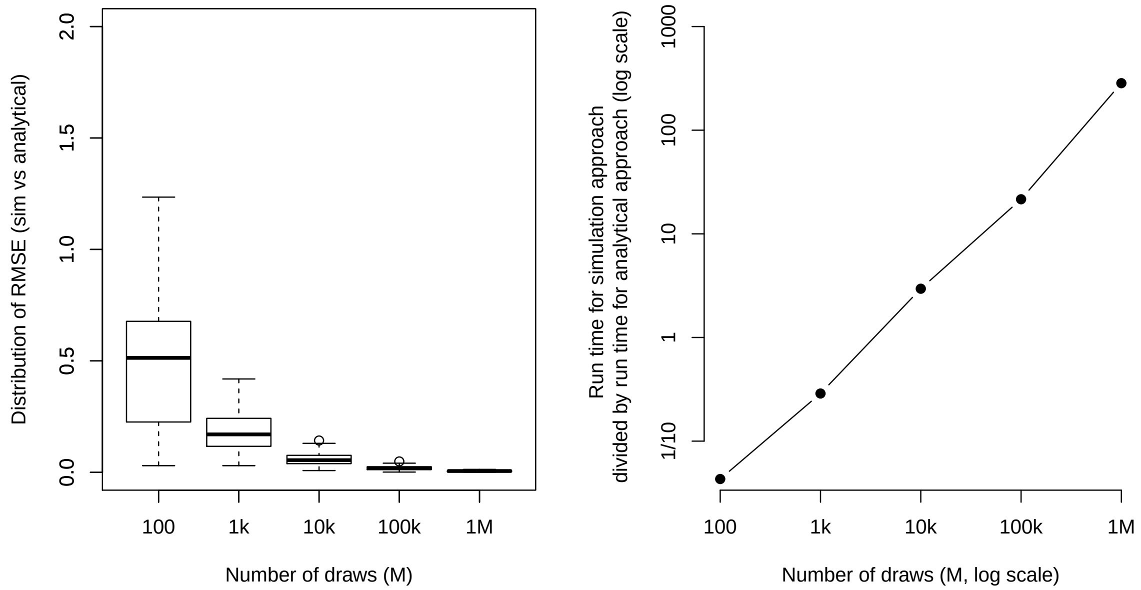
We begin by drawing J sets of Dirichlet parameter values at which we will calculate pivotal probabilities. Specifically, for scenario j we (1) draw a vector $\overline{\mathbf{v}}_j = \{\overline{v}_{ab,j}, \overline{v}_{ac,j}, \overline{v}_{ba,j}, \overline{v}_{bc,j}, \overline{v}_{ca,j}, \overline{v}_{cb,j}\}$ from a Dirichlet distribution with parameters $\{6, 4, 5, 5, 4, 6\}$ and (2) draw γ_j independently from a uniform distribution between 15 and 60. Together, $\overline{\mathbf{v}}_j$ and γ_j define beliefs for scenario j. For each of these J scenarios there are 12 pivotal probabilities to compute. Let $\mathbf{T}$ denote the $J \times 12$ matrix of pivotal probabilities computed with our numerical approach, and let $\tilde{\mathbf{T}}_M$ denote the $J \times 12$ matrix of pivotal probabilities computed with our simulation method using M draws from the belief distribution. Our focus is on how the discrepancies between $\mathbf{T}$ and $\mathbf{T}_M$ vary with M. We summarize these discrepancies with two approaches.

First, for each M and for each of $J = 100$ we compute the root mean squared error (RMSE), or average discrepancy, between $\mathbf{T}$ and $\tilde{\mathbf{T}}_M$. That is, for a given M, we compute the RMSE for each row of $\mathbf{T}$ and $\tilde{\mathbf{T}}_M$. The left panel of Figure 8 summarizes the distribution of these 100 RMSEs at each value of M. It shows that the distribution of RMSEs converges toward a point mass at zero as the number of draws from the belief distribution increases. As the simulation approach becomes more accurate, its computational burden also increases (as shown in the right

[32]Let p denote the probability of the pivotal event occurring (i.e. a finishes within $\frac{1}{2N}$ of b), and let p' denote the probability of the *expanded* pivotal event occurring (i.e. a finishes within $\delta = \frac{k}{N}$ of b). Let X denote the number of pivotal events observed in M trials and X' denote the number of expanded pivotal events observed in M trials. We propose to measure $\frac{X'}{kM}$ instead of $\frac{X}{M}$. If $p' \approx kp$ our measure will be approximately unbiased. But the variance will be lower (as p goes to zero) by a factor of k:

$$\begin{aligned}
\mathrm{Var}\left(\frac{X}{M}\right) &= \frac{1}{M^2}\mathrm{Var}(X) = \frac{Mp(1-p)}{M^2} = \frac{p(1-p)}{M} \\
\mathrm{Var}\left(\frac{X'}{kM}\right) &= \frac{1}{k^2M^2}\mathrm{Var}(X') = \frac{Mp'(1-p')}{k^2M^2} \approx \frac{Mkp(1-kp)}{k^2M^2} = \frac{p(1-kp)}{kM} \\
\lim_{p\to 0} \frac{\mathrm{Var}\left(\frac{X'}{kM}\right)}{\mathrm{Var}\left(\frac{X}{M}\right)} &\approx \lim_{p\to 0} \frac{\frac{p(1-kp)}{kM}}{\frac{p(1-p)}{M}} = \lim_{p\to 0} \frac{1-kp}{k(1-p)} = \frac{1}{k} \qquad (9)
\end{aligned}$$

Figure 7: Numerical/analytical approach agrees with simulations but is many times faster



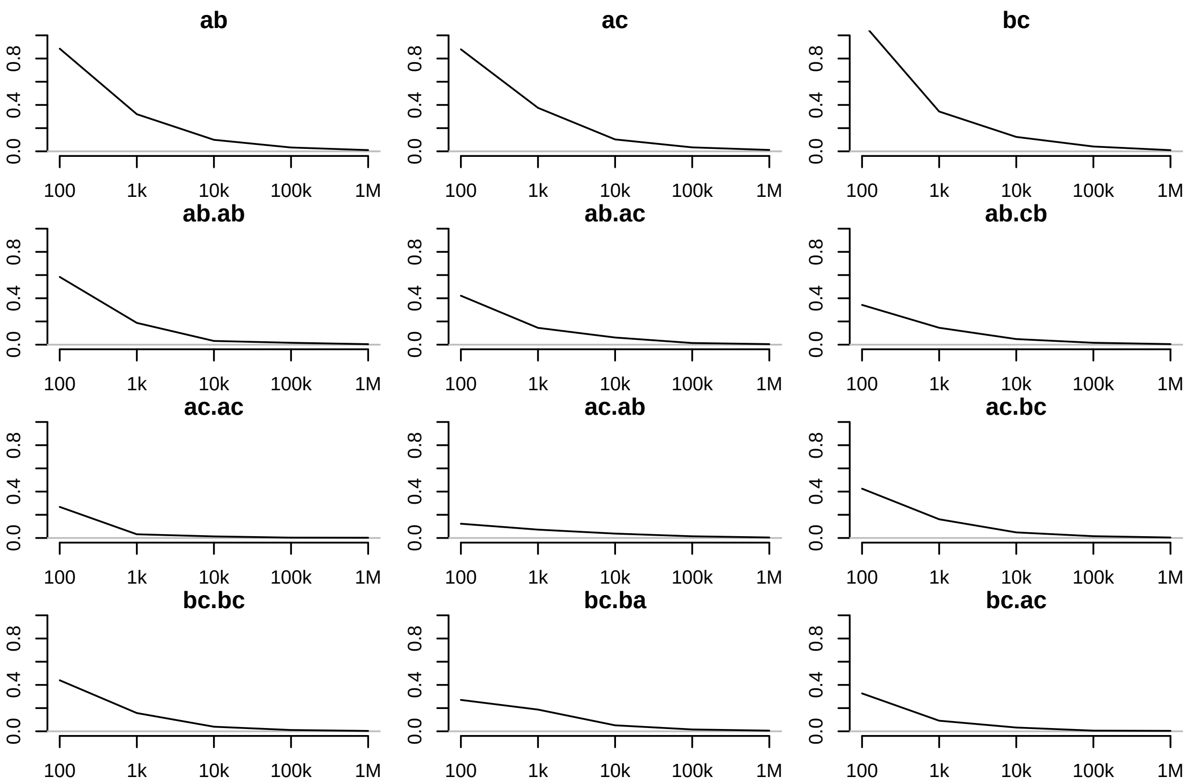
Note: For each of 100 sets of belief parameters, we compute pivotal probabilities (1) analytically and (2) by simulation, with M draws from the belief distribution. We then calculate the RMSE across the 12 pivotal events between the analytical approach and the simulation approach for each of the 100 scenarios. The left figure shows, for each value of M (horizontal axis), that the distribution of the RMSEs across the 100 scenarios converges to a point mass at zero as the number of simulation draws increases. The right panel shows how the relative computational burden of the simulation approach increases as the number of simulation draws increases

panel): with M of 1 million, our machine takes over 250 times longer to compute the pivotal probabilities by simulation than by the analytical approach.[33]

Second, for each pivotal event we compute at each M the RMSE across the $J = 100$ scenarios between $\mathbf{T}$ and $\tilde{\mathbf{T}}_M$. That is, for a given M, we compute the RMSE for each column of $\mathbf{T}$ and $\tilde{\mathbf{T}}_M$. Figure 8 summarizes how these RMSEs vary with M. It shows that the RMSE drops toward zero for all pivotal events as the number of number of draws from the belief distribution increases.

[33]Benchmarking performed on a 2017 MacBook Pro with 2.3 GHz processor and 16GB memory.

Figure 8: RMSE by pivotal event and number of draws in simulation



Note: For each of 100 sets of belief parameters, we compute pivotal probabilities (1) analytically and (2) by simulation, with M draws from the belief distribution. The figure shows, for each pivotal event, the average discrepancy (RMSE) between the two approaches as M increases.