Proportionality for ranked voting, in theory and practice

GERDUS BENADÈ, CHRISTOPHER DONNAY, MOON DUCHIN, and THOMAS WEIGHILL

Classical social choice theory includes a long list of criteria, or fairness axioms, for elections where individuals rank their preferences. Famous impossibility theorems from the 1970s concern the properties of voting rules to convert profiles of ranked preferences to winner sets. But though public perceptions of fairness are strongly keyed to proportional representation, notions of proportionality are strikingly missing from the standard roster of fairness axioms. We design a framework to measure *the degree of proportionality of seats to voter preference* under a wide class of systems for electing legislative bodies, even when elections are conducted without party labels. We begin by building out a set of generative models for creating synthetic ranked preference profiles, with an emphasis on flexibility and realism; in particular, we can efficiently generate polarized elections with properties motivated by the extensive body of work on racially polarized voting in the United States. The models use notions of *blocs* of voters and their *slates* of preferred candidates, which need not be known to voters but could be implicit in their voting patterns. The models serve as a thought tool for building a new definition of proportional representation and provide a framework that allows researchers to compare systems of election in terms of their tendency to produce proportional outcomes. We illustrate this by giving both empirical and theoretical results for single transferable vote (STV) elections.

This work brings a statistical modeling toolkit to the questions around ranked choice voting and proportionality. At the same time, it builds a much-needed bridge from computational social choice theory to political science, where degrees of proportionality have been intensely studied for well over a century, and to the work of practitioners in current reform efforts around voting rights and democracy.

ACM Reference Format:
Gerdus Benadè, Christopher Donnay, Moon Duchin, and Thomas Weighill. 2024. Proportionality for ranked voting, in theory and practice. 1, 1 (November 2024), 34 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

Authors' address: Gerdus Benadè; Christopher Donnay; Moon Duchin; Thomas Weighill.

ACM XXXX-XXXX/2024/11-ART
https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

In this paper, we give what we believe to be the first definition of *the degree of proportionality of votes to seats* that is general enough for use with ranked preferences and any structure of districts and voting rules that fills a legislative body.[1] This fills a gap in the classical social choice literature. Ken Arrow's foundational work studied social choice functions that combine multiple input rankings into one output ranking; following this, a series of important results were conjectured and proved from the 1960s to the 1990s concerning the use of rankings to output winner sets. Impossibility theorems of Müller–Satterthwaite, Gibbard–Satterthwaite, and Duggan–Schwartz rule out the viability for single-winner or multi-winner elections of simultaneously securing multiple axioms of fairness (see, for instance, [Taylor, 2002]). Examples of fairness axioms from early social choice theory include strategy-proofness, monotonicity, and the Condorcet criterion. However, these simply do not rank high in the public discourse around democracy.

Another area of need in the computational social choice literature is in defining generative models of election using domain knowledge of real-world electoral dynamics. We construct novel generative models of ranking that are inspired by polarized elections in real-world settings; in particular, voting rights law in the United States has used notions of voting blocs and their degrees of cohesiveness for decades. (The term "generative model" is often associated with large language models as paradigms of artificial intelligence; here, what is being generated is realistic voting rather than realistic language.) With these models and access to observed electoral data, we can test voting rules on both real and synthetic preference profiles, yielding information—some provable and analytic and some qualitative and simulation-based—on whether roughly proportional outcomes do indeed tend to result from so-called "semi-proportional" systems.

1.1 Contributions

New generative models. Generative models of voting use parameters and data—in our case, historical voting patterns, demographics, cohesion parameters, and candidate strength—to build a probability distribution from which ballots are sampled and elections can be simulated. In this paper we build and test generative models. These are the first mechanisms for producing ranked ballots that incorporate polarization according to candidate slates. We will offer some validation that our models comport far better with real-world ranking data than previous models (solid coalitions, IC, IAC), which builds our confidence in using them to analyze voting rules.

Rethinking proportionality. The proportionality of representation for a subgroup of voters could have a very simple interpretation in demographic terms (the group's seat share is in line with its share of the electorate). However, this fails to account for any complexity in the voting patterns of that group and the complementary voters. We define a framework that replaces demographic proportionality for a bloc of voters with *support proportionality* for a slate of candidates: the slate's seat share should be in line with the combined support for its candidates. We note that this kind of proportional representation is broader than that of PR systems such as party list voting, which secure support proportionality—on the basis of party only—by construction, so that the finding of proportional outcomes is vacuous on that axis. Here, we are measuring a kind of proportionality that is endogenous or emergent with respects to votes cast, and can be measured not only on the basis of party but with respect to any other cohesive preference. In other words, voters might not even be aware of which candidates constitute a slate; slates can be identified after the fact on the basis of trends in voter behavior.

[1] In particular, both a collection of single-winner elections and a collection of multi-winner elections are covered in our framework. All other notions we are aware of work by recourse to approval ballots, as we describe further below.

Incorporating domain knowledge. This project engages domain knowledge in voting rights law and practice in multiple ways. First, we shift the definition of voter cohesion to match the ordinary and legal use of the term. In the previous social choice literature, definitions of *cohesive* groups of voters tend to revolve around overlapping approval ballots: for instance, Sánchez-Fernández et al. [2017] call a group of voters ℓ–cohesive, where n candidates are running for k seats, if they comprise at least $\ell n/k$ people and their preferences overlap in at least ℓ candidates. This nuances earlier notions in which "cohesion" requires only a non-empty overlap in approvals. By contrast, this paper introduces notions of cohesiveness keyed to the probability of members of a group to support candidates from a certain slate. Compare this to, for instance, the landmark *Thornburg v. Gingles* decision of the U.S. Supreme Court, requiring Voting Rights Act plaintiffs to ascertain "whether members of a minority group constitute a politically cohesive unit" by measuring whether "a significant number of minority group members usually vote for the same candidates."[2] Expert work supporting a finding of cohesiveness revolves around "statistical evidence of voting patterns" using past elections, and polarization is typically summarized by using standard inference techniques to estimate the share of support for slates of candidates by blocs of voters [Hebert et al., 2010]. The authors of the present paper are drawing on just this kind of experience in voting rights expert work.[3]

Secondly, definitions related to justified representation are far from notions of proportionality in the political science literature and the popular vernacular: seat share in line with vote share. The relationship of seat share to vote share has been intensely studied at least since the late 19th century, and measurement of deviation from ideal seats/votes curves has generated a significant literature in the last fifty years especially in the work of Tufte, King, Grofman, and many more.

Finally, our use of ranked ballots rather than approval ballots is aligned with practice (and reform momentum) in the United States and internationally. Several U.S. states have recently debated adoption of ranked choice elections: Maine and Alaska now use ranked voting for statewide elections, with Nevada midway through the process of enacting a shift. Dozens of cities from San Francisco to Minneapolis use ranked choice for municipal elections, and New York City recently switched to ranked choice to elect city councillors and the mayor. Outside of the U.S., ranked choice voting is used for local or legislative elections in much of the Anglophone world—including Scotland, Ireland, New Zealand, and Australia—as well as for parliamentary elections in Malta and Papua New Guinea. As ranked voting is considered more broadly, stakeholders are increasingly asking about its properties, and one claim in common circulation is that they deliver more proportional outcomes for minority voting groups than could be expected from first-past-the-post systems. We seek to investigate these claims.

Illustrating with STV. While our notion of proportionality and the generative models we propose do not assume use of any specific voting rule, we will use *single transferable vote* (STV) as a test case. STV is a family of voting rules within ranked choice voting, using a transfer mechanism for selection of multiple winners, where the number of seats to be filled in a single contest is called the *magnitude*. In STV elections, there is a threshold level of support needed to be elected—typically the threshold is about $1/(k+1)$ of the first-place votes, where k is the magnitude. The election is conducted in rounds. As candidates are either elected (by passing the threshold) or eliminated from contention, the (surplus) votes supporting those candidates are transferred to the next options on

[2] *Thornburg v. Gingles* (1986), https://www.oyez.org/cases/1985/83-1968.
[3] For instance, consider recent expert work in Texas: minority racial groups were estimated to collectively support Democratic candidates in general elections from 2012–2020 at rates of 85-92%, while white voters supported Republican candidates at rates of 75-85% in the same contests. Expert report of REDACTED, *TX NAACP et al. v. Abbott*, Case No. 1:21-CV-00943-RP-JES-JVB.

their respective ballots.[4] We note that *instant runoff voting* or IRV, an extremely popular alternative in practice, is the same voting rule as STV in the special case $k = 1$.

Though STV is the basis for the examples in this paper, the express goal of the work is to set up a framework suitable for the comparative study of any voting rules applied to ranked ballots.[5]

1.2 Related work

Statistical ranking models, or models that assign a probability to permutations on a set of elements, have been studied at least since the early 20th century, going back to Thurstone [1927]. Subsequent models include those introduced by Bradley and Terry [1952], Plackett [1975], and Luce [1959], which form the basis for the BT and PL models in this paper, respectively. Benter [2008] introduced a variation of the Plackett model with a dampening parameter to account for less careful deliberation of lower-ranked items. Johnson et al. [2002] proposed a model to combine rankings that were determined by several different sources—which could have used different methods and criteria—into an aggregate, or meta, ranking scheme.

Ranking models have been used in a variety of applications in the broader social science literature. Stern [1990] applies the methods to horse races, where the marginal probability of each horse finishing first is known in advance. Bradlow and Fader [2001] apply time series models to Billboard "Hot 100" list, to show how song rankings change over time. Graves et al. [2003] apply a combination of ranking models to racecar competition outcomes. In the area of election analysis, Upton and Brook [1975] fit a Plackett model to ranked ballots in London elections to determine the effect of candidate name ordering on the ballots, also known as positional bias. Gormley and Murphy [2008] fit a combination of Plackett-Luce and Benter models to polling data from Irish elections in 1997 and 2002. In particular, they find the models to be effective in identifying voting blocs (groups of voters with similar ranked preferences) within the electorate. In the same paper, the authors fit mixtures of Plackett-Luce models to cast vote records from Irish elections, with the main goal of identifying blocs within the electorate.[6] These analyses are descriptive, based on historical data. In a recent paper, Garg et al. [2022] model outcomes of elections in multi-member Congressional districts under a solid coalition assumption, which means that the ballots are effectively unranked (and do not differentiate candidates within each coalition).

Our work is related in several respects to the existing computational social choice literature. There is a large body of work on the axiomatic properties of voting rules in various settings, including notions with a family resemblance to proportionality. The classical fairness axiom called Proportionality for Solid Coalitions (PSC), introduced by Dummett [1984], has been widely noted to be inadequate because it only applies with perfectly solid voting blocs, which never occurs in practice. The chief examples of axioms improving on PSC are those of (extended) justified representation (JR/EJR) [Aziz et al., 2017], which are structured as guarantees under approval-based multi-winner voting: sufficiently large groups whose approvals have non-trivial overlap can't be shut out of the winner set. Refer to Lackner and Skowron [2022] for a more thorough discussion. Various papers have used proportionality language for functions that map approval ballots to ranked outcomes [Skowron et al., 2017] and, quite recently, for functions that carry ranked ballots

[4]Specific mechanics vary; in this paper we have implemented the vote-tallying mechanism used by Cambridge, MA for its City Council elections, except as noted below.

[5]When single-winner rules like IRV are used to elect a representative body, as in the New York City Council, the framework here will be applicable with respect to similar candidates across the full candidate pool.

[6]In the language that will be introduced below, this roughly corresponds to fitting a Name-PL model (see Remark 4) with unknown group sizes and no slate structure. That is, their method is designed to learn preferences for all candidates by each of two blocs. Fitting a mixture model in this way does not produce a partition of candidates into slates so it is not clear how it might fit with a notion of proportionality.

to sets of approval ballots, and from there map to multi-winner outcomes [Brill and Peters, 2023]. While similar in spirit, it would be difficult to compare ideas invoking justified representation to ours directly because the JR family of axioms relies on a fundamentally different definition of cohesion. Furthermore, like PSC, these axioms are binary: a winner set satisfies rank-PJR+, for instance, or it does not. These definitions are not keyed to giving degrees of proportionality.

In terms of generative models of election, numerical experiments in this literature traditionally rely on assumptions of *impartial culture* [Pritchard and Wilson, 2009], under which voters are independent and every permutation of candidates is equally likely, *impartial anonymous culture*, in which Lebesgue measure is used to set relative preferences, or use *spatial* or distance-based models [Elkind et al., 2017, Tideman and Plassmann, 2010]. See Szufa et al. [2022, 2020] for a comparison of common generative models (called "statistical cultures") and a recent discussion of how to sample approval elections.

Spatial models [Enelow and Hinich, 1984] represent voters (and candidates) as ideal points in a metric space—in other words, using a space with a distance function as the latent space for voter preferences—and are common across fields. Voters are presumed to vote either deterministically for their closest representatives or probabilistically (upweighting closer candidates) [Burden, 1997]. Two commonly used methods for estimating ideal points (typically from Congressional roll-call data) are NOMINATE [Poole and Rosenthal, 1985] and IDEAL [Clinton et al., 2004]. Ranked choice voting models can be built from spatial models. For example, Gormley and Murphy [2007] combine a spatial and Plackett-Luce model to analyze Irish STV elections (discussed further in §5), and Kilgour et al. [2020] use a spatial model (where voters rank by proximity) to measure the effect of ballot truncation on single-winner ranked choice outcomes. Garg et al. [2022] also use a spatial model in one section, with voter ideal points extracted from ideology ratings in a commercial voter file, to relate the "diversity" of elected officials to the sizes of multimember districts.

Spatial models on one hand, and approval votes on the other, are favored by the mathematically inclined because they lend themselves to provable theoretical properties of voting rules. For example, under the implicit utilitarian voting framework, ordinal votes are proxies for underlying utilities and the *distortion* of a voting rule captures its worst-case loss compared to having full information [Procaccia and Rosenschein, 2006]. Anshelevich et al. [2018] study the distortion of STV under metric preferences, and Gkatzelis et al. [2020] recently settled a well-known conjecture on the optimal metric distortion when aggregating rankings to elect a single winner.

Our goal is to strike out in a new direction, with definitions that enable new questions to surface.

2 BLOCS, SLATES, AND PROPORTIONALITY

2.1 Defining blocs, slates, and notions of preference

The concept of blocs and slates is straightforward: *slates* are disjoint sets of candidates, such that voter propensity to support the various slates can be measured. The idea that voters would exhibit a preference among slates makes sense for an electorate overall, or when split out into disjoint groups of voters we call *blocs*.

To make this precise, we must delineate what it means for the preference profile consisting of ranked votes from a group of voters to display an overall preference for one group of candidates over another. We list several notions of preference or propensity that can be measured in an observed vote profile—that is, these are measurements that can be made on any cast vote record that has been minimally cleaned so that each ballot is a partial ranking (a permutation of a subset of the candidates).

Definition 2.1. Suppose an election is conducted with bloc structure $(A, \mathcal{A}, B, \mathcal{B})$ consisting of sets of voters A, B and corresponding slates of candidates $\mathcal{A} = \{A_1, \ldots, A_r\}$ and $\mathcal{B} = \{B_1, \ldots, B_s\}$.

This accommodates an election with one bloc (where we will adopt the convention $A = \emptyset$) or two, and this definition is easily expanded to more blocs. Often, we will use A for the majority bloc and B for the minority, when those are clear. The preference definitions are given below for bloc B.

Suppose voters are allowed to rank up to $n \le r + s$ candidates on their ballots—that is, ballots may be incomplete rankings of varying length, up to some maximum.

- Bloc B prefers slate $\mathcal{B}$ with *first-place preference* p_B if the share of first-place votes in the profile for $\mathcal{B}$ candidates is p_B.
- Bloc B prefers slate $\mathcal{B}$ with *positional preference* $P_B = (p_1, p_2, \ldots, p_n)$ if the share of ballots placing an $\mathcal{B}$ candidate in position i (among those for which a vote is cast and neither slate was exhausted in the higher positions) is p_i. In particular, the special case of *consistent positional preference* p_B corresponds to $P_B = (p_B, p_B, \ldots, p_B)$.
- Given a positional scoring rule with weights $(w_1, w_2, \ldots, w_n)$, we say that B prefers slate $\mathcal{B}$ with *score preference* p_B if the share of their score for $\mathcal{B}$ candidates is p_B. The default option will be to give standard Borda weights to the top k ranks via the score vector $(k, k-1, \ldots, 1, 0, \ldots, 0)$ in a magnitude-k election; we will refer to this as (top-k) *Borda preference.* For the purpose of Borda scoring, partial rankings are completed with an averaging convention (see §A).

Preferences for the A bloc are defined analogously; the only difficulty in extending to *more* than two blocs is one of cumbersome notation.

We will interpret each of these preference parameters as an indication of how *cohesive* bloc B is, with higher preference parameters (closer to 1) indicating more strongly aligned blocs.

Example 2.2. Suppose an election has been conducted with $r = 3, s = 2, n = 5$ (i.e., complete rankings are allowed), and suppose the voters are labeled as A voters or B voters. Suppose that the summarized preference profile for the B bloc is given by

$\times 2$	$\times 3$	$\times 8$	$\times 1$	$\times 5$	$\times 3$	$\times 5$
B_1	B_1	B_1	A_1	B_2	B_2	B_1
B_2	A_2	B_2	B_1	B_1	A_3	B_2
A_1	B_2	A_2	B_2	A_1	A_1	
A_2	A_3	A_1		A_3	B_2	
A_3	A_2			A_2	A_2	

(by name)

i.e.,

$\times 7$	$\times 3$	$\times 8$	$\times 1$	$\times 3$	$\times 5$
B	B	B	A	B	B
B	A	B	B	A	B
A	B	A	B	A	
A	A	A		B	
A	A			A	

(by slate)

Then the first-place preference of the B bloc for $\mathcal{B}$ candidates is $26/27$, the positional preference is $(\frac{26}{27}, \frac{21}{27}, \frac{4}{7}, \frac{3}{3}, -)$, the Borda preference to all five places is $232/405$ with ballot completion, and the top-2 Borda preference is $73/81$. Note that the last few positional scores are $4/7$, $3/3$, and undefined—rather than $4/22$, $3/21$, and 0—because of only considering ballots which have not exhausted the B candidates.

2.2 Defining proportionality

If the electorate is undivided ($A = \emptyset$) and the voters support slate $\mathcal{B}$ with propensity π_B, then we interpret that as voters giving the slate π_B share of their support. In this case, the proportionality ideal is extraordinarily simple: seat share equals vote share, i.e.,

$$S_B = \pi_B.$$

When voters only select a single candidate, this is exactly the vernacular notion of proportionality.

When there are two distinct blocs with different voting behavior that partition the whole electorate, this extends by convex combination to a natural heuristic for a proportional outcome of an election. If π_B is the preference parameter for bloc B towards its candidates and likewise π_A for bloc A, then the proportionality target sets seat share S_B for the $\mathcal{B}$ slate at

$$S_B = N_B \cdot \pi_B + (1 - N_B)(1 - \pi_A),$$

where N_B is the share of voters from the B bloc. That is, proportionality pins the representation to combined support for $\mathcal{B}$ candidates: the size of the B bloc times its level of cohesion (the propensity to vote for $\mathcal{B}$ candidates) plus the size of the complementary bloc times its level of crossover voting (again, the propensity to vote for $\mathcal{B}$ candidates).[7]

This enables us to say, for instance, whether a particular election outcome was near-proportional (in a given bloc structure, if applicable) with respect to first-place preferences, or to Borda preferences, or any other notion of propensity. Proportionality is not a foregone conclusion for ranked choice voting even in the extremely simple case where the blocs are defined by first-place votes; lower-ranked choices may or may not track closely with first-place preference.

Example 2.3. We use a sample of nine real-world Scottish local government STV elections to illustrate how to use the definition in practice. We illustrate with both first-place preference and score preference.

		first-place pref.		top-k Borda share		STV outcome
election	(r, s, k)	π_B	proportionality	π_B	proportionality	
North Ayrshire 2022 Ward 1	$(9, 3, 5)$	0.35	1.75 seats	0.35	1.75 seats	2 seats
Angus 2012 Ward 8	$(4, 2, 4)$	0.39	1.56 seats	0.40	1.60 seats	2 seats
Clackmannanshire 2012 Ward 2	$(5, 3, 4)$	0.49	1.96 seats	0.53	2.12 seats	2 seats
Aberdeen 2022 Ward 12	$(7, 3, 4)$	0.48	1.92 seats	0.48	1.92 seats	2 seats
Aberdeen 2017 Ward 12	$(7, 3, 4)$	0.36	1.44 seats	0.39	1.66 seats	2 seats
Falkirk 2017 Ward 6	$(3, 3, 4)$	0.56	2.24 seats	0.57	2.28 seats	2 seats
Renfrewshire 2017 Ward 1	$(5, 3, 4)$	0.43	1.72 seats	0.42	1.68 seats	2 seats
Fife 2022 Ward 21	$(5, 3, 4)$	0.45	1.80 seats	0.45	1.80 seats	2 seats
Glasgow 2012 Ward 16	$(9, 3, 4)$	0.41	1.64 seats	0.39	1.56 seats	2 seats

Table 1. Here, s is the number of $\mathcal{B}$ candidates (defined by membership in the Scottish National Party and the Greens), r is the number of candidates from all other parties, and k is the number of seats to be filled in the election. We treat the electorate as a single bloc (undivided) and measure π_B as the level of first-place support for the $\mathcal{B}$ slate, and the share of top-k Borda scores, respectively, for the $\mathcal{B}$ slate.

We adopt a simplified slate structure where two Scottish parties—Scottish National and Green—are defined as a slate $\mathcal{B}$, and the complementary slate $\mathcal{A}$ combines all other parties.[8] In Table 1 we consider the level of proportionality in two ways. We first use first-place preference to define π_B, the propensity of voters to support slate $\mathcal{B}$. (Equivalently, we can think of this as defining blocs by first-place vote and adopting 100% cohesion.) This means that the number of seats needed to

[7]One could consider alternative definitions of proportionality, for example, based on a bloc-weighted combination of the number of seats a slate wins in each of the hypothetical elections in which only one of the blocs participates. However, this requires fixing a voting rule. We deliberately propose a notion of proportionality that is agnostic to the choice of voting rule.
[8]The other parties include Conservatives, Labour, Liberal Democrats, multiple parties defined by their stance on independence from the UK, far-right parties like the National Front, and some farther-left socialist parties. STV can be tested with respect to any slate, though securing proportionality by first-place preference will be most likely when the slate has cohesive support from a subset of voters.

achieve (first-place) proportionality is $\pi_B \cdot k$, the proportional seat share times the number of seats. Applying an alternative choice of propensity, we can use π_B to measure the slate-$\mathcal{B}$ share of the top-k Borda scores to arrive at a different proportionality target.

3 GENERATIVE MODELS

3.1 Constructing the models

In this section, we set up generative models of election, including several variants derived from classical statistical ranking literature in the style of Plackett-Luce and Bradley-Terry models.[9] All five models produce rankings of candidates; one of the five includes partial rankings at a rate keyed to historical data. The statistical ranking models will be introduced with what we call Slate vs. Name versions: the Slate versions begin by constructing an abstract ballot type before filling in candidate names, while the Name versions work directly with candidate names. Though at first the by-name and by-slate versions may seem extremely similar, we find that Slate-PL and Slate-BT have several desirable properties compared to Name-PL and Name-BT. These, together with a model called the Cambridge sampler (Slate-CS), make up the generative models explored in the empirical work in this paper. In this paper we focus on settings with 1-2 blocs for notational convenience but the framework immediately expands to more groups.

Definition 3.1. For all of the models below, assume a fixed bloc structure $(A, \mathcal{A}, B, \mathcal{B})$ with $\mathcal{A} = (A_1, \ldots, A_r)$ and $\mathcal{B} = (B_1, \ldots, B_s)$, allowing the possibility that $A = \emptyset$ as before.

A *ballot* is a partial or complete ranking of the $r + s$ candidates and a *ballot type* is a partial or complete permutation of the symbols $A^r B^s$, i.e., a simplified ballot that treats the candidates of each slate as indistinguishable from each other.

The models below will use the following parameters to generate a profile for bloc B:

Cohesion Tendency of the bloc to support slate $\mathcal{B}$, given as a parameter $\pi_B \leq 1$ (typically required to be at least $1/2$ in the multi-bloc case).

Strength Tendency of bloc B to agree on preferred candidates *within* each slate. This consists of probability vectors $I_{BA} = (a_1, \ldots, a_r)$ and $I_{BB} = (b_1, \ldots, b_s)$; i.e., the entries are non-negative and sum to one. For instance, if $I_{BA} = (.1, .8, .1)$, then typical B voters strongly prefer candidate A_2 to A_1 or A_3.

We can combine the cohesion and strength data into a single probability vector

$$I_B = \big((1 - \pi_B)a_1, \ldots, (1 - \pi_B)a_r, \quad (\pi_B)b_1, \ldots, (\pi_B)b_s\big).$$

Using these components, we can define five generative models as follows. The first two work directly with ballots, while the latter three first construct ballot types. These are analogous to the profile by name and the profile by slate in Example 2.2.

Name-PL Plackett-Luce by name: Each B-bloc voter chooses candidate i to be ranked first with probability $I_B(i)$. They continue to select candidates for lower-ranked positions in order, at each stage selecting candidate j with probability proportional to $I_B(j)$. In other words, each voter samples their ballot without replacement from all candidates proportional to their weighting in I_B.

Name-BT Bradley-Terry by name: The probability that a B voter casts a ballot σ is proportional to

$$\prod_{i <_\sigma j} \frac{I_B(i)}{I_B(i) + I_B(j)},$$

[9]Earlier versions of the Name-PL, Name-BT, and Slate-CS models have been discussed in unpublished work by an overlapping collection of authors. References are suppressed here for anonymization purposes.

where $i <_\sigma j$ means that i is ranked before (i.e., higher than) j in σ. In other words, for each pairwise comparison of candidates, we introduce a term for the likelihood of ranking one before the other coming from the relative weights in I_B.

Slate-PL Plackett-Luce by slate: Each B-bloc voter chooses between the symbol A and B in the ith position with probability π_B of choosing B, as long as both $\mathcal{A}$ candidates and $\mathcal{B}$ candidates remain available. Once a slate is exhausted, the rest of the complete ranking is filled in with the remaining symbol.

Slate-BT Bradley-Terry by slate: Suppose a ballot type σ is a permutation of $A^r B^s$, that is, an ordered list containing r A symbols and s B symbols. Suppose that out of the rs comparisons of the instances of A with the instances of B, the A occurs earlier than the B a total of $0 \le i \le rs$ times. The probability that a B voter casts this ballot is proportional to $(1 - \pi_B)^i (\pi_B)^{rs-i}$.

Slate-CS Cambridge sampler: We draw from a dataset consisting of ten years of ranked votes from city council elections in Cambridge, MA. Historical candidates have been labeled as white (W) or as people of color (C), with help from local organizers. To use this model, we make a choice to designate bloc B as corresponding to voters who put a W candidate first ($B = W$), or who put a C candidate first ($B = C$). We use the cohesion parameter π_B to decide probabilistically whether the voter chooses their own slate or the other slate in the first position. Then we complete the ballot type by drawing with weight proportional to frequency from the cast ballots with that header.

In all three Slate models, we must then assign candidate names to the symbols A and B. We do so by drawing without replacement (Plackett-Luce style) from I_{BA} and I_{BB} separately to order $\mathcal{A}$ and $\mathcal{B}$, then fill in names accordingly.

Since a PL voter can be thought to fill in their ballot from top to bottom according to pre-computed preferences, we can think of this as modeling an "impulsive" voter. By contrast, a BT voter makes comparisons of every two entries on their ballot and weighs that ballot against one with some reversals, modeling a "deliberative" voter. These give new generative models to study, greatly expanding on the generative models in the literature, and they do so in a manner that comports well with U.S. voting rights law; we can plug in standard cohesion parameters for majority and minority groups as the π_A, π_B. We will give a brief validation showing model performance in matching a few observed elections in §3.2.2.

Remark 1 (PL preferences). *Slate-PL with $(A, \mathcal{A}, B, \mathcal{B})$ and any cohesion and candidate strength parameters is expected to produce blocs with consistent positional preference π_B (respectively π_A) for their own slates, and therefore with first-place preference π_B (or π_A) as well.*

Remark 2 (Names versus slates). *It turns out to be an important distinction to work directly with the names or to create a type first, then add names. The reason for the divergence is that the Slate models handle I_{BA} and I_{BB} separately; concatenating them into I_B before making length comparisons yields unintended results, such as a highly cohesive bloc whose voters tend to put their strong candidate first and then immediately cross over to supporting the opposite slate. These effects can be explored in the supplementary plots (§??) which compare all five models.*

Remark 3 (About the Cambridge data). *Cambridge, Massachusetts uses STV for its city council and school board elections and has done so since 1941. Our source of Cambridge historical data is city council elections to fill $k = 9$ seats by STV from 2009 to 2017, coded by candidate race as described above; there are frequently 20 or more candidates who run in each contest. If a ballot type is selected from the historical frequency histogram that has more candidates from a given slate than the (r, s) chosen for a given simulation run allows, then we ignore further instances. For instance, a ballot type of AAABB in an election where $r = s = 2$ will be read as AABB.*

One valuable aspect of our use of Cambridge historical data in the present study is that it lets us incorporate realistic short-ballot voting behavior without a proliferation of extra parameters. For instance, Cambridge voters cast "bullet votes" (listing only one candidate and leaving other positions blank) 7501 times out of 87,914 ballots cast in our data set, and this will be reflected in the ballots generated by the CS model. However, a serious limitation is that we have coded the candidates by race, while Cambridge city council politics are likely more polarized by other candidate features—for instance, an explicit slate of affordable housing candidates is routinely advertised before election day and is highly salient to voter behavior. Nevertheless, race is a candidate feature often apparent to voters which allows us to observe naturalistic patterns of alternation in voting.

Remark 4 (Mixture models). *The definitions above are in terms of specified blocs of voters with different voting preferences. However, there is a strong connection to mixture models suggested by the structure here. In a mixture model, each voter is assigned independently to a class, and then randomly submits a ballot based on the parameters for that class. More precisely, if N_1 and N_2 are the weights for two different classes of voter with $N_1 + N_2 = 1$, and μ_1 and μ_2 are two distributions on ballots corresponding to the two classes, the probability of a ballot σ is*

$$\mu(\sigma) = N_1\mu_1(\sigma) + N_2\mu_2(\sigma).$$

As the number of voters increases, the fraction of voters assigned to each class converges to N_1 and N_2 respectively; for large numbers of voters we can therefore consider the size of each class to be predetermined and treat voters as if they belong to two blocs of fixed size.

In particular, since it considers pairwise probabilities, the BT model with two blocs resembles a mixture of Mallows models. It differs in allowing swaps to be weighted by preference between slates rather than by their position in the ranking.

3.2 Visualization

3.2.1 MDS plot of vote profiles. One difficulty in studying ranked choice elections is that, unlike oversimplified Example 2.2, real-world elections frequently have too many valid ballots possible to effectively see the full preference profile. For instance, an election with six candidates can be thought of as having 1236 possible ballots to cast—there are 6! complete rankings and a roughly equal number of partial rankings.[10] Thinking of profiles as distributions over valid ballots allows us to define natural notions of distance between profiles, such as the L^1 distance between profiles given by the sum over possible ballots of the absolute value of the difference of shares for that ballot. (Up to a constant factor, this is the same as the total variation distance of distributions.) With this notion we can visualize differences between the generative models as we vary parameters.

To illustrate the importance of candidate strength, we introduce four out of the infinitely many variations on I_B concerning the preferences of B-bloc voters.

- **U** (uniform-uniform): preferences are uniform over $\mathcal{A}$ candidates and uniform over $\mathcal{B}$ candidates.
- **S** (strong-strong): preferences are strong over both slates, namely with some candidates receiving more support.
- **X** (uniform-strong): uniform support for $\mathcal{A}$ candidates and strong support for some $\mathcal{B}$ candidates;
- **Y** (strong-uniform): the reverse.

[10]Here, we identify a ballot of length 5 with a complete ranking of length 6, since the last-place candidate is implicit.

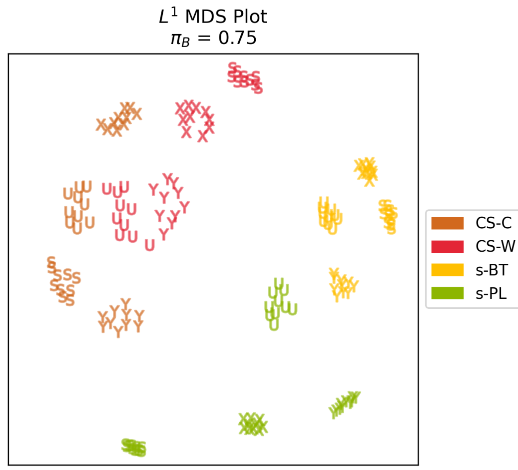

Fig. 1. Multi-dimensional scaling (MDS) plot for one-bloc profiles with $r = s = 3$ (3 candidates per slate), under a variety of generative models and candidate strength scenarios. Each model is designated by a different color, and the candidate strength scenarios are denoted U, S, X, Y, as described in the text. The pairwise distances between profiles are computed with L^1 distance on the distributions. Each preference profile has 1000 ballots, and we have generated 10 profiles by each of the 16 model/strength pairs. Note: it is not surprising that CS profiles, whether bloc B is identified with W-led or C-led ballots, fall far from PL and BT profiles, because PL and BT always generate complete rankings, while CS uses real historical data that includes many partial rankings. This observation can be used to give a sense of scale for the distances in the plot.

In the multi-dimensional scaling (MDS) plot in Figure 1, the first-place preference for $\mathcal{B}$ candidates is $\pi_B = .75$; Supplemental Figure 8 shows how the outputs change as π_B varies. In this plot, we can see some systematic differences and similarities.[11]

For instance, strength scenarios Y and X interpolate between U and S, as we might have expected. Also, BT profiles resemble both kinds of Cambridge outputs more than PL profiles do, though the reason for this is far less clear. (Compare Supplemental Figures 10–18, which bear this out from another point of view.)

3.2.2 Validation on Scottish elections. A benefit of using parameterized generative models is the possibility of fitting to real-world elections. Though we leave a full-bore fitting effort to future work, this section shows the potential of this approach to match the observed non-solidity of coalitions.

To this end, we define a *swap distance* between two ballot types, partial or complete. For complete ballots, this counts the smallest number of swaps of neighboring symbols necessary to transform one ballot type into the other; for instance, $\mathrm{dist}(AABBB, ABBAB) = 2$. See §A for a discussion of efficiently measuring this distance, including an extension to partial or weakly ranked ballots.

Using swap distance, we can investigate the extent to which vote profiles deviate from the solid coalition assumption. Let us return to the nine Scottish elections and the slate $\mathcal{B}$ discussed above. For every ballot cast in the election, we can compute its distance to the solid A-over-B ballot type A^sB^r. (Note that a solid vote of the opposite kind looks like B^rA^s, lying at distance rs.) For the

[11]The reader should recall that MDS plots are simply low-distortion planar embeddings, which depend on a choice of random seed. The x and y axes have no meaning; only the relative pairwise distances are meaningful with respect to the data. We have verified that the structure of the plots stays the same for a few choices of random seed.

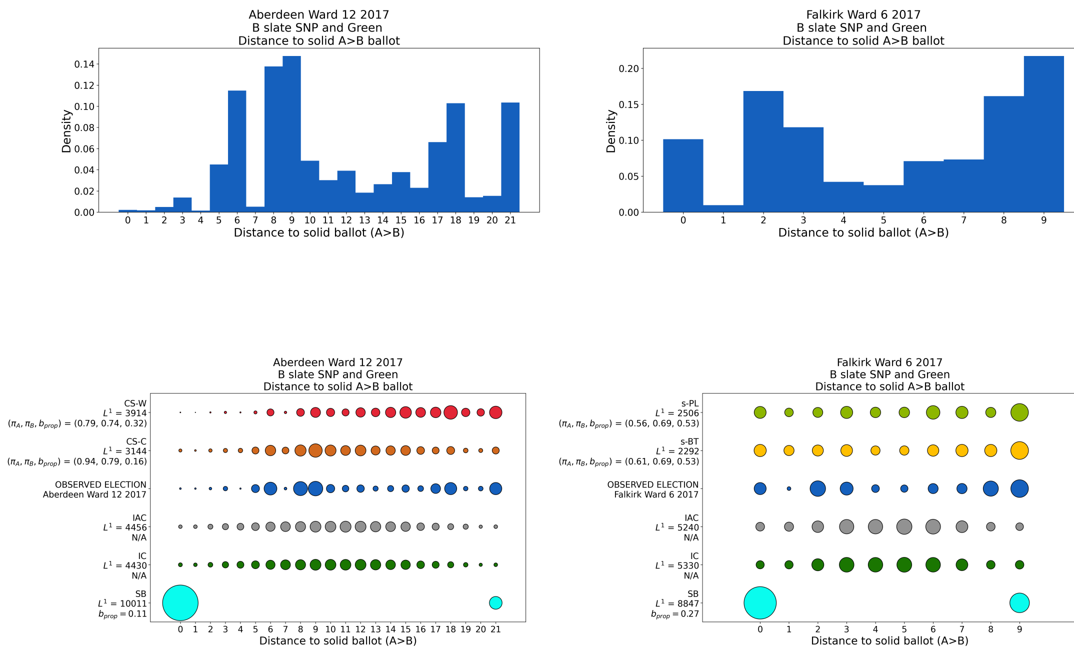

Fig. 2. Top: Histograms showing the distribution of swap distances to solid A-over-B type in Aberdeen Ward 12 and Falkirk Ward 6, 2017. Bottom: Bubble plots showing the distribution of swap distances, where the area of each circle is proportional to frequency. The top two colored rows show outputs from models introduced in this paper, with parameters optimized to match the observed election. The third row in dark blue is the observed election, for which the data exactly repeats the conventional histograms. The bottom three rows show the best fit when voters are constrained to solid ballots, followed by samples under IC (impartial culture) and IAC (impartial anonymous culture) models already popular in the social choice literature.

Aberdeen Ward 12 and Falkirk Ward 6 elections from 2017, these distances are summarized in the histograms of Figure 2.

Next, we can attempt to generate profiles that are the best match for these histograms using the models in §3.1. We can accomplish interesting results *even with an undivided electorate* (one bloc). We choose our cohesion parameter by optimizing π_B to minimize the L^1 distance to the observed election. The resulting distance distributions are visualized in the bubble plots of Figure 2 (and see Appendix C for a full range of outputs). The traditional assumption of solid coalitions produces distributions that are point masses at distances 0 and rs, which clearly have little in common with the real-world ballot distributions. Both visually and in terms of measured L^1 distance, the models do well at matching observed patterns of non-solidity of coalitions.

3.2.3 Parameter interactions. Next, we leverage the generative models in combination with a voting rule to produce simulations that highlight complex interactive effects between model parameters.

We vary N_B over $\{.05, .15, \ldots, .95\}$ and we vary both π_A and π_B over $\{.55, .65, .75, .85, .95\}$. We have selected four candidate strength scenarios for two blocs (based on the one-bloc scenarios in §3.2.1); these are chosen to give a small window on how powerfully candidate strength can interact with other factors.

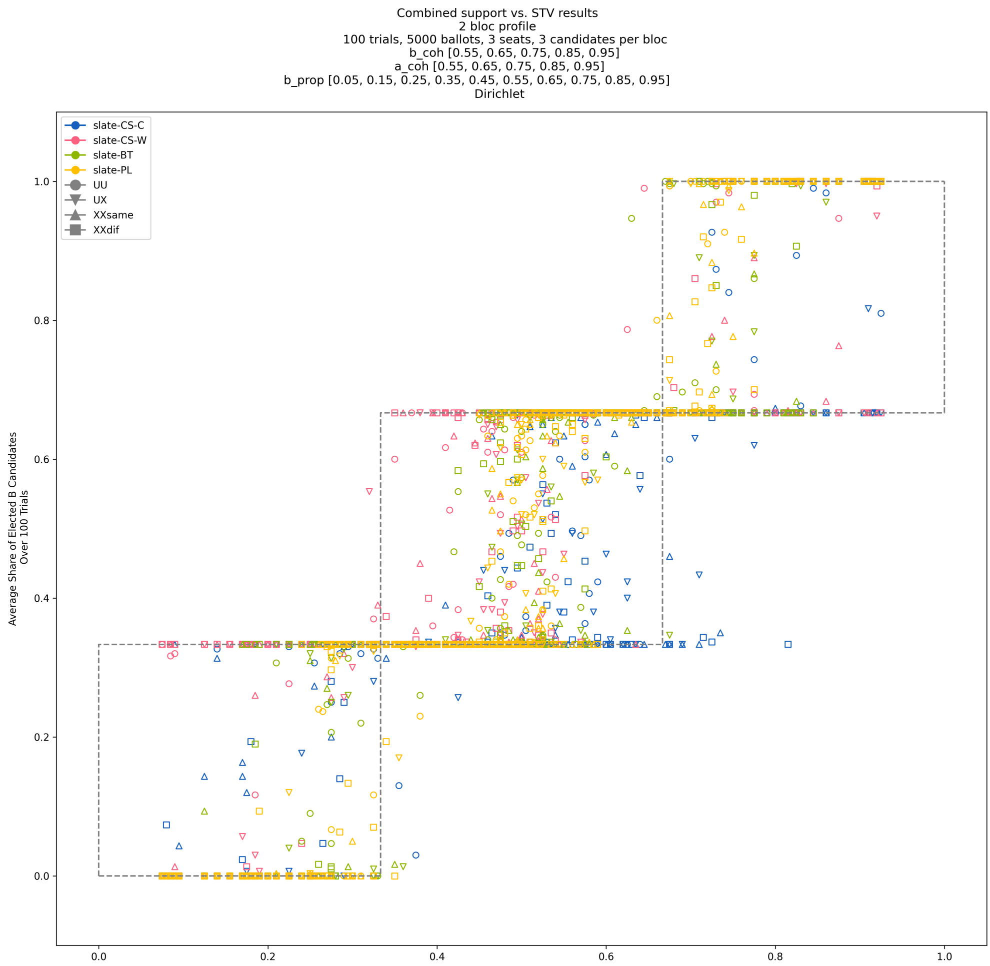

Fig. 3. Setting $(r, s, k) = (3, 3, 3)$, we independently vary the B proportion of the electorate, the generative model, the A and B cohesion, and the candidate strength settings. In this visualization, we have run 100 trials for each parameter tuple, recording the number of $\mathcal{B}$ candidates elected for each simulated profile. The x axis position is the combined support for $\mathcal{B}$ (with respect to first-place votes) and the y-axis position is the average number of seats over the trials with each tuple of parameters. The dashed lines show the proportionality target rounded up and down to the nearest whole number of seats.

- **UU** both blocs have uniformly random preference order over each slate;
- **UX**: I_{BB} has a strong candidate while others are uniform;
- **XX-same**: A and B blocs strongly prefer the same $\mathcal{B}$ candidate and are otherwise uniform;
- **XX-diff**: A and B blocs strongly prefer different $\mathcal{B}$ candidates and are otherwise uniform.

In effect, we must make five choices for each batch of runs: model, strength scenario, population share, cohesion for A voters, and cohesion for B voters. We then generate a batch of 100 profiles from each 5-tuple to place each symbol on the plot. The x-axis position is the combined support level for $\mathcal{B}$ candidates observed in the profiles, given by $N_B \cdot \pi_B + (1 - N_B)(1 - \pi_A)$ as above—so a given support level can be achieved in many different ways. The y-axis position is the average number of seats won by $\mathcal{B}$ candidates when the batch of profiles is run through the STV voting rule.

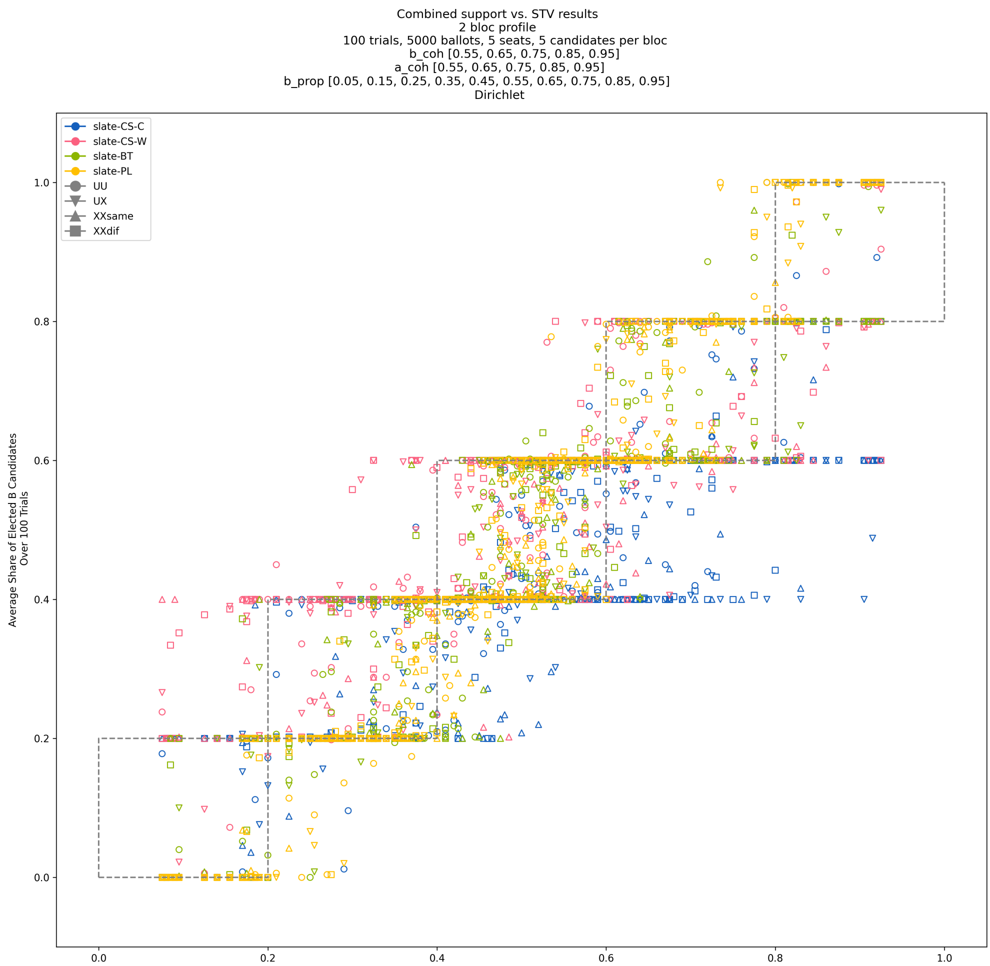

Fig. 4. This time $(r, s, k) = (5, 5, 5)$. We again independently vary the B proportion of the electorate, the generative model, the A and B cohesion, and the candidate strength settings. In this visualization, we have run 100 trials for each parameter tuple, recording the number of $\mathcal{B}$ candidates elected for each simulated profile. The x axis position is the combined support for $\mathcal{B}$ and the y-axis position is the average number of seats over the trials with each tuple of parameters. The dashed lines show the proportionality target rounded up and down to the nearest whole number of seats.

If the proportionality ideal were hit exactly, the symbols would all fall on the main diagonal. The proportionality target rounded up and down to whole numbers of seats is shown with dotted lines in the plots. For the most part, the data points fall within these proportionality targets.

4 ASYMPTOTIC PROPERTIES

In this section, we give proof of concept that the framework presented here is robust enough to admit provable statements about STV, a system of election for which theorems have so far been elusive. The standard assumption of solid coalitions, in particular, has every voter rank all candidates from one slate above all candidates from the other. This assumes away any role for transfer between slates. Therefore, although the assumptions below are strong, they are hugely more flexible than what has existed in the literature so far. In particular, the results of §4.1 fix the candidate order within slates, but allow probabilistic crossover between slates.

The first aim of the results in this section is to show that the generative models let us formulate and prove nontrivial quantitative results; beyond that, we can change one feature of the voting system at a time and see how results differ, as in Proposition 4.1 vs. Proposition 4.2, where a small change in how STV is tabulated can make a large difference. Corollary 4.4 is a surprisingly strong numerical bound on the seats-to-votes ratio for large elections.

4.1 Single bloc asymptotics

In this section, we focus on the case of one bloc of voters and two slates of candidates. Note that even with a single bloc the fact that we have two slates means any lack of cohesion immediately leads to the richer types of crossover ballots that motivated our generative models.

For the Slate-PL and Name-PL models, we can prove theoretical results that offer a kind of asymptotic generalization of the well-known Proportionality for Solid Coalitions (PSC). We give asymptotics as the number of voters goes to infinity, since our models are probabilistic.

We start by giving bounds on the outcomes for a bloc voting under Slate-PL model. The results reveal that the choice of precise method for tallying votes has a profound impact on the expected outcomes. With that in mind, we define two different methods for deciding which candidates are elected in each round of an STV vote tallying process.

- **Simultaneous election:** if multiple candidates exceed the threshold for election in a certain round, they are all elected and their excess votes transfer down to the remaining candidates before the next round.
- **One-by-one election:** if multiple candidates exceed the threshold for election in a certain round, the one with the most votes is elected and their excess votes are transferred. The tallying process then proceeds to the next round.

Based on the way that election results are reported by the city of Cambridge, it appears that Cambridge follows the simultaneous election method.[12]

Proposition 4.1 (Slate-PL, fixed order, simultaneous election). *Consider an STV contest with simultaneous election for k open seats, a single bloc of N voters, and two slates of candidates $\mathcal{A}$ and $\mathcal{B}$. Suppose that the voters vote according to a Slate-PL model and all voters rank the candidates within each slate in a fixed order, $A_1 > A_2 > \cdots > A_k$ and $B_1 > B_2 > \cdots > B_k$. (Further candidates would therefore be irrelevant.) Write $\alpha = \pi_A$ and $\beta = 1 - \pi_A$ for the tendency to support $\mathcal{A}$ and $\mathcal{B}$ candidates, respectively, and assume $\frac{1}{2} < \alpha < 1$. Then the share of $\mathcal{B}$ candidates elected satisfies*

$$\frac{1}{2} - \frac{1}{2k}\left(\frac{\alpha}{\beta} - 1\right) \leq S_B \leq \frac{1}{2} \qquad \text{a.a.s. as } N \to \infty.$$

Thus, the share of $\mathcal{B}$ candidates elected approaches $1/2$ with high probability as k, N get large, even if their support β is very small.

[12]See for instance https://www.cambridgema.gov/Election2023/Official/Council%20Round.htm

Proof. In this proof, we will treat the ballot shares as deterministically equaling their expectations, so that any strict inequalities we derive stay true with high probability as $N \to \infty$. First, observe that the fixed order means that in a given round, only one A candidate and one B candidate has any first-place votes. The use of the Droop quota means that $N/(k+1)$ votes is the threshold of election; when a candidate reaches that threshold, they are elected and their excess votes are transferred, leading to a reduction of the mass of total votes by $N/(k+1)$; ballots can not be exhausted because Slate-PL produces complete rankings. Until the full k seats are elected, this means that there is always enough vote mass to elect at least one candidate per round, and since only two have first-place support, either one candidate is elected, or one A and one B are elected simultaneously. Suppose that one candidate from each slate is elected in rounds $1, \ldots, \ell$, but that only one candidate is elected in round $\ell + 1$; this must be an $\mathcal{A}$ candidate with high probability, because $\alpha > \beta$. So in round $\ell + 1$, the support for $\mathcal{B}$ has dropped below threshold. Solving for ℓ, we have

$$\left(1 - \frac{2\ell}{k+1}\right)\beta \geq \frac{1}{k+1}; \quad \left(1 - \frac{2(\ell+1)}{k+1}\right)\beta < \frac{1}{k+1},$$

and this gives

$$(k - 2\ell - 1)\beta < 1 \leq (k - 2\ell + 1)\beta \implies k - \frac{\alpha}{\beta} - 1 \leq 2\ell < k - \frac{\alpha}{\beta}.$$

Thus there are at least $\frac{k}{2} - \frac{\alpha}{2\beta} - \frac{1}{2}$ candidates from $\mathcal{B}$ elected, as claimed.

For the upper bound, note that the only round in which a single B can be elected immediately follows a round in which a single A was elected. This is because after one of each is elected (whether this occurs simultaneously or in successive rounds), all first-place votes have been cleared and the full mass of remaining ballots now favors the next A candidate with a share close to α. □

Figure 5 gives a visualization of Proposition 4.1. To obtain the exact $N \to \infty$ asymptotics plotted in the figure, we allow a fractional number of ballots of each kind, and assume that the number of ballots of each kind is exactly equal to the expectation under the model. We also assume that vote transfers are fractional and deterministic.

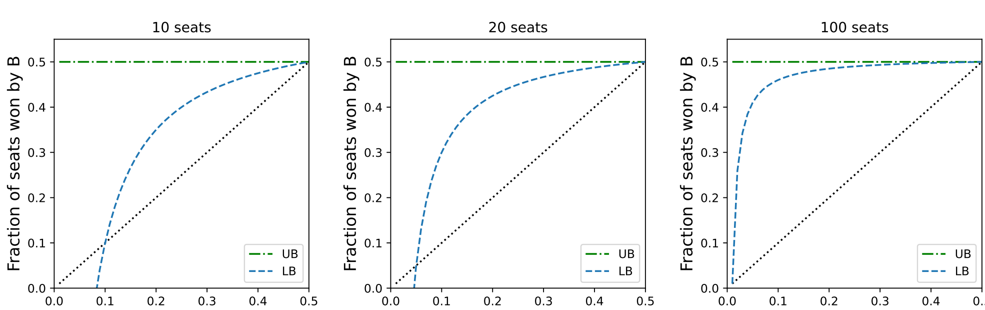

Fig. 5. A visualization of the lower bound and upper bounds in Proposition 4.1 for various values of β.

It is somewhat surprising that, as $k \to \infty$, $\mathcal{A}$ and $\mathcal{B}$ are equally represented even though all voters prefer $\mathcal{A}$. Proposition 4.1 assumes simultaneous election transfers—this, together with the fact that there are fixed rankings over $\mathcal{A}, \mathcal{B}$, creates a situation where in nearly every round all first-place votes land on the top remaining $\mathcal{A}$ and $\mathcal{B}$ candidates, and both are elected.

We now consider the one-by-one vote tallying method. A practical difference between the simultaneous and one-by-one elections is that one-by-one election may exhibit a kind of leap-frogging, where a candidate who is over the threshold in round 1 may nonetheless be elected after a candidate who was below the threshold in round 1. This does not happen in simultaneous elections.

Proposition 4.2 (Slate-PL, fixed order, one-by-one election). *Take the same setup assumptions as in the previous proposition except for using one-by-one election rather than simultaneous election. Assume that* $\log_\alpha(1/2)$ *is not an integer. Let* $\gamma = \lfloor \log_\alpha(1/2)) \rfloor$, *so that* $\alpha^\gamma \geq 1/2$ *but* $\alpha^{\gamma+1} < 1/2$. *Then seat share* S_B *for B candidates satisfies*

$$\frac{\frac{k+2}{\gamma+2}-1}{k} \leq S_B \leq \frac{1}{2} - \frac{\delta(k+1)}{k}\left(\frac{\lceil \frac{2\alpha-1}{\frac{t}{1-\ell t}-2\alpha\ln\alpha}\rceil}{1+\lceil \frac{2\alpha-1}{\frac{t}{1-\ell t}-2\alpha\ln\alpha}\rceil} - \frac{1}{2}\right) \quad \text{a.a.s. as } N \to \infty.$$

for any $\delta \in [0,1]$. *By setting* $\delta = \frac{\sqrt{k}-1}{\sqrt{k}}$ *we obtain that as* $k \to \infty$, *the value* $\hat{S}_B$ *to which* S_B *tends a.a.s. satisfies*

$$\frac{1}{\gamma+2} \leq \hat{S}_B \leq \frac{1}{1+\lceil \frac{2\alpha-1}{-2\alpha\ln\alpha}\rceil}$$

Proof. Consider the tally after ℓ rounds. Let the share of all ballots (live or not) currently headed by an A be $\hat{\alpha} \leq 1-\ell t$. In the next round, the share of ballots headed by an A is $(\hat{\alpha}-t)\alpha$. After round $\ell+s-1$, assuming only A candidates have been elected since round ℓ, the share of ballots headed by an A is $\hat{\alpha}\alpha^{s-1} - t\alpha^{s-1} - \cdots - t\alpha$. Since the total mass of ballots left at that stage is $1-(s-1+\ell)t$, it follows that the difference between the share headed by an A and the share headed by a B is given by

$$\Delta(s) = 2\hat{\alpha}\alpha^{s-1} - 1 - t(2\alpha^{s-1} + \cdots + 2\alpha) + t(s-1+\ell) \tag{1}$$

$$= 2\hat{\alpha}\alpha^{s-1} - (1-\ell t) - t\left(\frac{2\alpha(1-\alpha^{s-1})}{1-\alpha} - (s-1)\right) \tag{2}$$

In particular, an A will be elected next if and only if $\Delta(s)$ is positive.

We define a *sequence* as a set of consecutive rounds consisting of electing A candidates, followed by a B candidate. If $\ell = 0$, then $\hat{\alpha} = \alpha$. By Lemma 4.3, $\Delta(s) \leq 2\alpha^s - 1$ for all $s \leq \gamma$, so $\Delta(s)$ is negative for $s = \gamma + 1$. Thus the first sequence has length at most γ.

If round ℓ was the first round of some later sequence, then all ballots headed by an A transferred last round, so $\hat{\alpha} \leq \alpha(1-\ell t)$. Thus

$$\Delta(s) \leq (1-\ell t)(2\alpha^s - 1) - t\left(\frac{2\alpha(1-\alpha^{s-1})}{1-\alpha} - (s-1)\right)$$

and using Lemma 4.3 again, we have that this is negative for $s = \gamma + 1$. Thus this sequence has length at most $\gamma + 1$, since we allow for round ℓ to be the first round of the sequence.

Suppose there are r sequences, followed possibly by electing a final set of A candidates. The best case for A candidates is if this final set consists of $\gamma + 1$ A candidates. In that case there are $\gamma + (r-1)(\gamma+1) + \gamma + 1 = r(\gamma+1) + \gamma$ A candidates elected and r B candidates elected. Since $k = r(\gamma+1) + \gamma + r = (r+1)(\gamma+2) - 2$, we obtain

$$S_A \leq \frac{k-r}{k} = \frac{k - \frac{k+2}{\gamma+2} + 1}{k}$$

We now turn our attention to the lower bound. Suppose that we are in round ℓ at the start of a sequence. By Lemma 4.3 and the fact that $\hat{\alpha} \geq (1-\ell t)\alpha$, we have

$$\begin{aligned}\Delta(s) &\geq 2(1-\ell t)\alpha^s - (1-\ell t) - t(2(s-1)-(s-1))\\ &\geq 2(1-\ell t)\left(\alpha s\ln(\alpha) + (\alpha - \alpha\ln\alpha)\right) - (1-\ell t) - t(s-1).\end{aligned}$$

This last expression is decreasing in s. The root of the expression is given by

$$s = 1 + \frac{2\alpha - 1}{\frac{t}{1-\ell t} - 2\alpha\ln\alpha}$$

Defining

$$\lambda(\ell) := \lceil \frac{2\alpha - 1}{\frac{t}{1-\ell t} - 2\alpha\ln\alpha} \rceil$$

we have that $\lambda(\ell)$ is a lower bound on the number of A candidates elected, starting at round ℓ, before a B is elected.

Fix $\delta \in [0,1]$. Since a candidate is elected every round, there are k rounds total. Suppose that m is the number of rounds contained in sequences whose first round is before or equal to round $\lfloor \delta(k+1) \rfloor$, so that in particular, $m \geq \lceil \delta(k+1) \rceil$. The fraction of these m rounds in which an A is elected is at least $\lambda(\delta(k+1))/(\lambda(\delta(k+1))+1)$. Since an A is elected in round m, an A is elected in at least half of the remaining $k-m$ rounds. Thus the seat share for A satisfies

$$\begin{aligned}S_A &\geq \frac{1}{k}\left(m \cdot \frac{\lceil \frac{2\alpha-1}{\frac{t}{1-\ell t}-2\alpha\ln\alpha} \rceil}{1 + \lceil \frac{2\alpha-1}{\frac{t}{1-\ell t}-2\alpha\ln\alpha} \rceil} + \frac{k-m}{2}\right)\\ &\geq \frac{1}{k}\left(\delta(k+1) \cdot \frac{\lceil \frac{2\alpha-1}{\frac{t}{1-\ell t}-2\alpha\ln\alpha} \rceil}{1 + \lceil \frac{2\alpha-1}{\frac{t}{1-\ell t}-2\alpha\ln\alpha} \rceil} + \frac{k-\delta(k+1)}{2}\right)\\ &= \frac{1}{2} + \frac{\delta(k+1)}{k}\left(\frac{\lceil \frac{2\alpha-1}{\frac{t}{1-\ell t}-2\alpha\ln\alpha} \rceil}{1 + \lceil \frac{2\alpha-1}{\frac{t}{1-\ell t}-2\alpha\ln\alpha} \rceil} - \frac{1}{2}\right)\end{aligned}$$

□

Lemma 4.3. *Let $\alpha \in [0.5, 1)$ and $s \in \mathbb{N}$.*

(a) $2\alpha^s \geq 2\alpha s\ln\alpha + 2(\alpha - \alpha\ln\alpha)$.

(b) *If* $1 \leq s \leq \log_\alpha(1/2) + 1$, *then*

$$\frac{\alpha(1-\alpha^{s-1})}{1-\alpha} \leq s - 1 \leq \frac{2\alpha(1-\alpha^{s-1})}{1-\alpha}$$

Proof. (a) Follows by considering the tangent line to $f(s) = 2\alpha^s$ at $s = 1$. (b) For the left hand bound, notice that since $\alpha < 1$,

$$\frac{\alpha(1-\alpha^{s-1})}{1-\alpha} < (1 + \alpha + \ldots + \alpha^{s-2}) \leq (s-1)$$

For the right hand bound, we have

$$\frac{2\alpha(1-\alpha^{s-1})}{1-\alpha} - (s-1) \geq \frac{2\alpha - \alpha^s}{1-\alpha} - \log_\alpha(1/2) \geq \frac{\alpha}{1-\alpha} - \log_\alpha(1/2)$$

The last expression is zero for $\alpha = 1/2$ and a derivative test shows that it is positive for all $\alpha \in (0.5, 1)$. □

Figure 6 contains a visualization of Proposition 4.2 using the same method to compute exact asymptotics as in Figure 5.

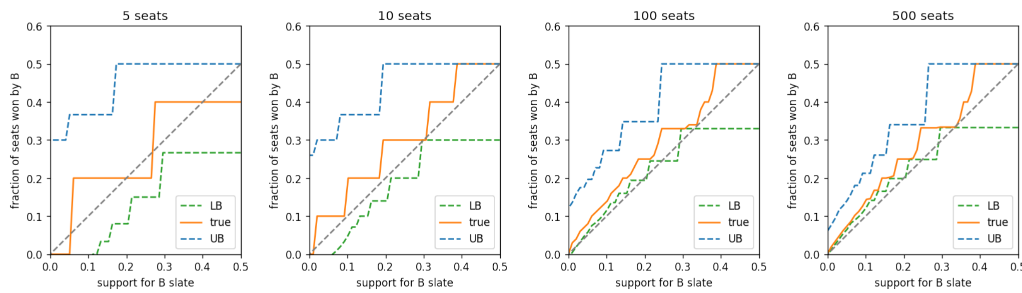

Fig. 6. Visualizations of the lower and upper bounds given by in Proposition 4.2 for $k = 5, 10, 100, 500$ and $\delta = (\sqrt{k} - 1)/\sqrt{k}$. The dotted line is $1 - \pi_A$, which is also B's combined share since there is no bloc B.

The preceding propositions let us assess the extent to which STV is likely to yield proportional representation in large elections, if the voters all adhere to a rigid ordering of candidates.

Corollary 4.4 (Bounding disproportionality for STV with fixed candidate orders). *Suppose we consider STV under the same conditions as above (slate-PL, fixed candidate order, sufficiently large k). Then under simultaneous election, disproportionality can get arbitrarily severe as the election gets large. However, under one-at-a-time election, the asymptotic ratio of seats to votes for the minority party satisfies*

$$\frac{2}{3} \le \frac{S_B}{\beta} \le 2,$$

where β is the support for B candidates.

Proof. Simultaneous election means that $\mathcal{B}$ candidates will tend toward 1/2 seat share, no matter their level of support from voters.

Working with the asymptotic value for $N \to \infty$ for one-by-one election, we have, as $k \to \infty$,

$$S_B/\beta \ge \frac{1}{(1-\alpha)\left(\frac{\ln(1/2)}{\ln \alpha} + 2\right)}$$

which is an increasing function of $\alpha \in [0.5, 1)$ and at $\alpha = 0.5$ achieves a minimum of $2/3$. For the upper bound, we have

$$S_B/\beta \le \frac{1}{(1-\alpha)\left(1 + \frac{2\alpha - 1}{-2\alpha \ln \alpha}\right)}$$

which has a supremum of 2 for $\alpha \in [0.5, 1)$.

□

Finally, when one bloc votes by Name-PL, asymptotic results are easy to describe for extreme candidate strength scenarios, assuming there are more candidates in each slate than seats open, and equal numbers of candidates in each slate.

Proposition 4.5 (Name-PL). *For ballots generated by a Name-PL model, the STV winners (with either vote tallying process) are (a.a.s.) the top candidates by support value, up to a choice about how to break ties between equally supported candidates. Thus we obtain the following results a.a.s. as $N \to \infty$. Without loss of generality, assume the preference for $\mathcal{A}$ candidates is $\alpha \ge 1/2$.*

(a) *If there are strong preferences within slates, so that $\alpha a_1 > (1-\alpha)b_1 > \alpha a_2 > (1-\alpha)b_2 > \cdots$, then the number of $\mathcal{A}$ and $\mathcal{B}$ candidates elected is equal or within one, no matter the value of α.*
(b) *On the other hand, if the support is divided uniformly within slates and $\alpha > 1/2$, then only A candidates are elected.*

Proof. To prove the first statement, consider a Plackett-Luce model with probability vector $(c_1, \ldots, c_k)$. For any partial ranking of candidates $\sigma' = (C_1 > C_2 > \ldots > C_\ell)$, meaning C_1 is the first choice and so on, let $F(\sigma', i)$ be the proportion of ballots which begin with $C_1 > C_2 > \ldots > C_\ell > C_i$. Asymptotically almost surely, if $i, j \notin \sigma'$, we have $c_i < c_j \implies F(\sigma', i) < F(\sigma', j)$. It follows that, initially, the candidates with the most first-place votes are (a.a.s.) those with the highest support values. Moreover, after vote transfers, the candidates with the most first places will be (a.a.s.) those remaining candidates (not yet elected or eliminated) with the highest support values. This proves the main statement, and (a) and (b) follow. □

4.2 Two-bloc asymptotics with fixed candidate order

We conclude our consideration of electoral outcomes with an observation that the asymptotics of two-bloc elections for the one-by-one variant of STV interpolate between solid coalitions and unpolarized voting in an intuitive way. (See Figure 7.)

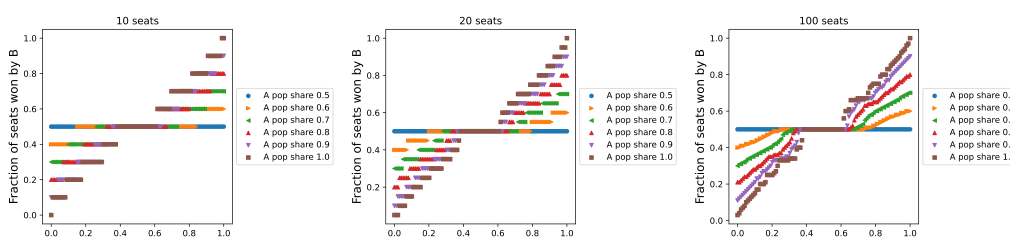

Fig. 7. Exact asymptotics (as the number of voters gets large) showing the share of seats won by the A bloc as their vote share and cohesion varies. The elections have $m = 10, 20, 100$ seats, with an inexhaustible supply of candidates. We use the Slate-PL model, suppose both blocs use the same fixed ordering over $\mathcal{A}$ and $\mathcal{B}$ and apply the one-by-one election variant of STV defined in §4.

One interesting (and real) artifact visible in these plots is that the outcome with seat share of 50% is a plateau that occurs for a range of cohesion values. To get an idea of the reason for this, note that since this plot assumes both blocs use a fixed candidate order $A_1, A_2, \ldots$ and $B_1, B_2, \ldots$, the first candidate elected with $\pi_A, N_A > .5$ will always be A_1. For large numbers of seats, where the election threshold is close to zero, there is a phase transition when $\pi^2 = (1-\pi) + \pi(1-\pi)$, occurring at $\pi = 1/\sqrt{2} \approx .707$, that determines whether the first transfer results in the election of A_2. For smaller π, enough support will transfer to B_1 that they are next to be elected. Similar polynomial thresholds determine how many A candidates are elected between successive B candidates. For π approaching 1/2, the order of election will alternate $ABABAB\ldots$, giving 1/2 seat share to each side.

5 CONCLUSION AND FUTURE WORK

In §3.2.2 we make first steps toward fitting models and parameters to realistic elections, with immediate payoff in a starkly improved correspondence to Scottish ranked elections than solid coalitions could offer. A more comprehensive fitting effort along these lines—simultaneously learning optimal blocs and slates from observed elections—is a natural future project. This would also point the way to new methods of measuring the degree of polarization, which can feed back usefully into voting rights law.

Our goal in this paper is to lay the groundwork to systematically study the tendency of systems to deliver more or less proportional outcomes for voters. Crucially, the framework we propose allows but does not require party labels, so that we can also identify emergent blocs with similar voting behavior after an election has been conducted. Finally, the new generative models outlined here can be theoretically explored, opening up rich directions for mathematical study, but can also give decision-makers a powerful toolkit for practical electoral reform.

REFERENCES

Elliot Anshelevich, Onkar Bhardwaj, Edith Elkind, John Postl, and Piotr Skowron. 2018. Approximating optimal social choice under metric preferences. *Artificial Intelligence* 264 (2018), 27–51.

Haris Aziz, Markus Brill, Vincent Conitzer, Edith Elkind, Rupert Freeman, and Toby Walsh. 2017. Justified representation in approval-based committee voting. *Social Choice and Welfare* 48, 2 (2017), 461–485.

William Benter. 2008. Computer based horse race handicapping and wagering systems: a report. In *Efficiency of racetrack betting markets*. World Scientific, 183–198.

Ralph Allan Bradley and Milton E. Terry. 1952. Rank Analysis of Incomplete Block Designs: I. The Method of Paired Comparisons. *Biometrika* 39, 3/4 (1952), 324–345. http://www.jstor.org/stable/2334029

Eric T Bradlow and Peter S Fader. 2001. A Bayesian Lifetime Model for the "Hot 100" Billboard Songs. *J. Amer. Statist. Assoc.* 96, 454 (2001), 368–381. https://doi.org/10.1198/016214501753168091

Markus Brill and Jannik Peters. 2023. Robust and Verifiable Proportionality Axioms for Multiwinner Voting. In *Proceedings of the 24th ACM Conference on Economics and Computation*. 301–301.

Barry C. Burden. 1997. Deterministic and Probabilistic Voting Models. *American Journal of Political Science* 41, 4 (1997), 1150–1169. http://www.jstor.org/stable/2960485

Joshua Clinton, Simon Jackman, and Douglas Rivers. 2004. The statistical analysis of roll call data. *American Political Science Review* (2004), 355–370.

Moon Duchin and Kris Tapp. 2024. Ballot Clustering Algorithms. *preprint* (2024).

Michael Dummett. 1984. Voting procedures. (1984).

Edith Elkind, Piotr Faliszewski, Jean-François Laslier, Piotr Skowron, Arkadii Slinko, and Nimrod Talmon. 2017. What do multiwinner voting rules do? An experiment over the two-dimensional euclidean domain. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 31.

James M Enelow and Melvin J Hinich. 1984. *The spatial theory of voting: An introduction.* CUP Archive.

Nikhil Garg, Wes Gurnee, David Rothschild, and David Shmoys. 2022. Combatting gerrymandering with social choice: The design of multi-member districts. In *Proceedings of the 23rd ACM Conference on Economics and Computation*. 560–561.

Vasilis Gkatzelis, Daniel Halpern, and Nisarg Shah. 2020. Resolving the optimal metric distortion conjecture. In *2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*. IEEE, 1427–1438.

Isobel Claire Gormley and Thomas Brendan Murphy. 2007. A Latent Space Model for Rank Data. In *Statistical Network Analysis: Models, Issues, and New Directions*, Edoardo Airoldi, David M. Blei, Stephen E. Fienberg, Anna Goldenberg, Eric P. Xing, and Alice X. Zheng (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 90–102.

Isobel Claire Gormley and Thomas Brendan Murphy. 2008. Exploring voting blocs within the Irish electorate: A mixture modeling approach. *J. Amer. Statist. Assoc.* 103, 483 (2008), 1014–1027.

Todd Graves, C Shane Reese, and Mark Fitzgerald. 2003. Hierarchical models for permutations: Analysis of auto racing results. *J. Amer. Statist. Assoc.* 98, 462 (2003), 282–291.

J Gerald Hebert, Paul Smith, Martina Vandenburg, and Michael DeSanctis. 2010. The Realist's Guide to Redistricting: Avoiding the Legal Pitfalls. American Bar Association.

Valen E Johnson, Robert O Deaner, and Carel P Van Schaik. 2002. Bayesian analysis of rank data with application to primate intelligence experiments. *J. Amer. Statist. Assoc.* 97, 457 (2002), 8–17.

D Marc Kilgour, Jean-Charles Grégoire, and Angèle M Foley. 2020. The prevalence and consequences of ballot truncation in ranked-choice elections. *Public Choice* 184 (2020), 197–218.

Martin Lackner and Piotr Skowron. 2022. Approval-based committee voting. In *Multi-Winner Voting with Approval Preferences*. Springer.

R Duncan Luce. 1959. Individual choice behavior. (1959).

Robin L Plackett. 1975. The analysis of permutations. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 24, 2 (1975), 193–202.

Keith T Poole and Howard Rosenthal. 1985. A spatial model for legislative roll call analysis. *American Journal of Political Science* (1985), 357–384.

Geoffrey Pritchard and Mark C Wilson. 2009. Asymptotics of the minimum manipulating coalition size for positional voting rules under impartial culture behaviour. *Mathematical Social Sciences* 58, 1 (2009), 35–57.

Ariel D Procaccia and Jeffrey S Rosenschein. 2006. The distortion of cardinal preferences in voting. In *International Workshop on Cooperative Information Agents*. Springer, 317–331.

Luis Sánchez-Fernández, Edith Elkind, Martin Lackner, Norberto Fernández, Jesús Fisteus, Pablo Basanta Val, and Piotr Skowron. 2017. Proportional justified representation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 31.

P Skowron, M Lackner, M Brill, D Peters, and E Elkind. 2017. Proportional rankings. In *International Joint Conference on Artificial Intelligence (IJCAI 2017)*. Association for the Advancement of Artificial Intelligence.

Hal Stern. 1990. Models for distributions on permutations. *J. Amer. Statist. Assoc.* 85, 410 (1990), 558–564.

Stanisław Szufa, Piotr Faliszewski, Łukasz Janeczko, Martin Lackner, Arkadii Slinko, Krzysztof Sornat, and Nimrod Talmon. 2022. How to Sample Approval Elections? *arXiv preprint arXiv:2207.01140* (2022).

Stanisław Szufa, Piotr Faliszewski, Piotr Skowron, Arkadii Slinko, and Nimrod Talmon. 2020. Drawing a map of elections in the space of statistical cultures. In *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems*. 1341–1349.

Alan D. Taylor. 2002. The Manipulability of Voting Systems. *The American Mathematical Monthly* 109, 4 (2002), 321–337. http://www.jstor.org/stable/2695497

Louis L Thurstone. 1927. A law of comparative judgment. *Psychological review* 34, 4 (1927), 273.

T Nicolaus Tideman and Florenz Plassmann. 2010. The structure of the election-generating universe. (2010). Manuscript.

GJG Upton and D Brook. 1975. The determination of the optimum position on a ballot paper. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 24, 3 (1975), 279–287.

A SWAP DISTANCE AND BALLOT COMPLETION

The distance between two (complete) ballots measures the complexity of swaps to turn one ballot into the other. We will generalize this to ballot types (where candidates within each slate are indistinguishable from one another). We will allow swaps of individual neighboring candidates at unit cost as a special case of general individual swaps, whose cost is the difference in their positions. For instance, though ballot *ABC* could be transformed to *CBA* with three neighbor-swaps, its total cost will be just 2 because A and C can be exchanged directly, leaving B in place. To calculate this efficiently, we adapt a lemma from a preprint of Duchin and Tapp [2024]. Given an ordering of candidates, let the score vector sc of a ballot be defined as the vector of Borda scores earned by each candidate, so for the candidate order A, B, C we have

$$\mathrm{sc}(ABC) = (3, 2, 1); \quad \mathrm{sc}(CBA) = (1, 2, 3).$$

This admits a natural generalization to score vectors for incomplete ballots and weak orders over candidates; unmentioned candidates are regarded as being tied at the end of the ballot, and ties are handled as averages. There is also a merge/unmerge move for ballots with ties: merging or unmerging two sets in neighboring ballot positions costs one-half of the product of the set sizes. (Neighbor swaps are a special case realized by one merge and one unmerge.)

If the types are $\mathcal{A}$ and $\mathcal{B}$, let the score vector by type, denoted $\mathrm{sc}^{A|B}$, report each candidate's score as the average over its type, so that for an election with $(r, s) = (2, 3)$ (i.e., $\mathcal{A} = \{A_1, A_2\}$ and $\mathcal{B} = \{B_1, B_2, B_3\}$), we have

$$\mathrm{sc}^{A|B}(AABBB) = (9/2, 9/2 \mid 2, 2, 2), \quad \mathrm{sc}^{A|B}(\{A, A\}, \{B, B, B\}) = (9/2, 9/2 \mid 2, 2, 2),$$

$$\mathrm{sc}^{A|B}(ABBAB) = (7/2, 7/2 \mid 8/3, 8/3, 8/3), \quad \mathrm{sc}^{A|B}(AB\{A, B, B\}) = (7/2, 7/2 \mid 8/3, 8/3, 8/3).$$

Here, the first two are sorted with $\mathcal{A}$ candidates before $\mathcal{B}$ candidates, and either of the second two can be sorted by moves incurring swap distance 2.

Lemma A.1 (Duchin–Tapp). *Swap distance on ballots can be calculated as an L^1 vector difference, as can the distance of a ballot type to being sorted. For ballots b_1, b_2,*

$$dist(b_1, b_2) = \frac{1}{2} \left\| sc(b_1) - sc(b_2) \right\|_1 ;$$

for a ballot type σ,

$$dist\,(\sigma, (\mathcal{A}, \mathcal{B})) = \frac{1}{2} \left\| sc^{A|B}(\sigma_1) - sc^{A|B}(\mathcal{A}, \mathcal{B}) \right\|_1 .$$

This is the distance to sorted that was used as a summary statistic of elections in §3.2.2.

B MORE MDS PLOTS

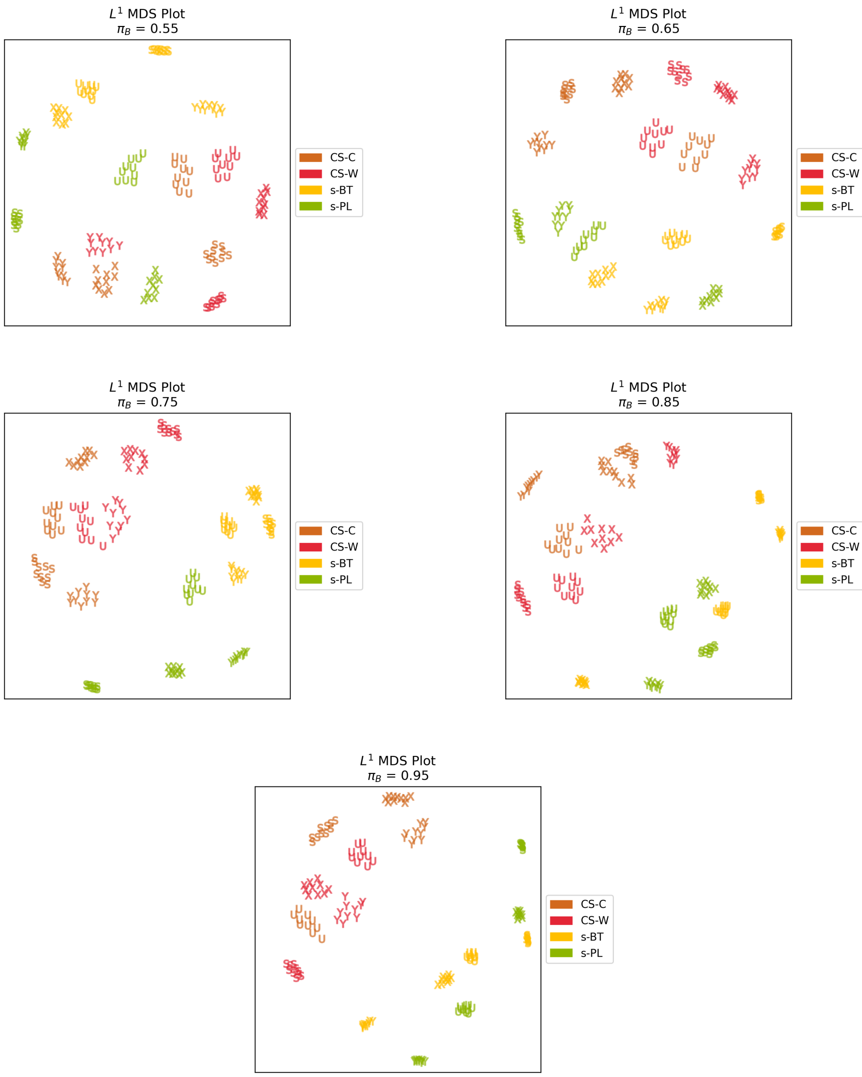

Fig. 8. Multi-dimensional scaling (MDS) plots for profiles with $r = s = 3$ (3 candidates per bloc), under a variety of generative models and candidate strength scenarios. Each model is designated by a different color, and the candidate strength scenarios are denoted U, S, X, Y, as described above. The pairwise distances between profiles are computed with L^1 distance on the profiles. Each preference profile has 1000 ballots, and we have generated 10 profiles by each of the 16 model/strength pairs. As $\pi_B \to 1$, the main difference appearing in the models is that the BT and PL profiles become tightly clustered for each candidate strength scenario, while the CS profiles remain more variable.

C FITTING TO SCOTTISH ELECTIONS

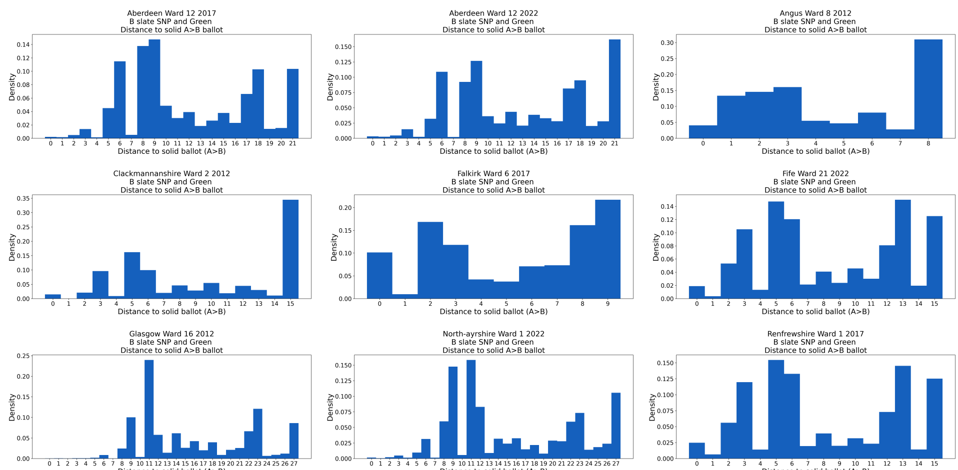

Fig. 9. Histograms showing the distribution of swap distances to solid A-over-B ballots in nine Scottish elections.

To conclude, we provide a full sweep of fitting outputs across the nine elections and various models in this paper. All simulations use the same number of ballots as in the observed election. Plots for all elections and models follow.

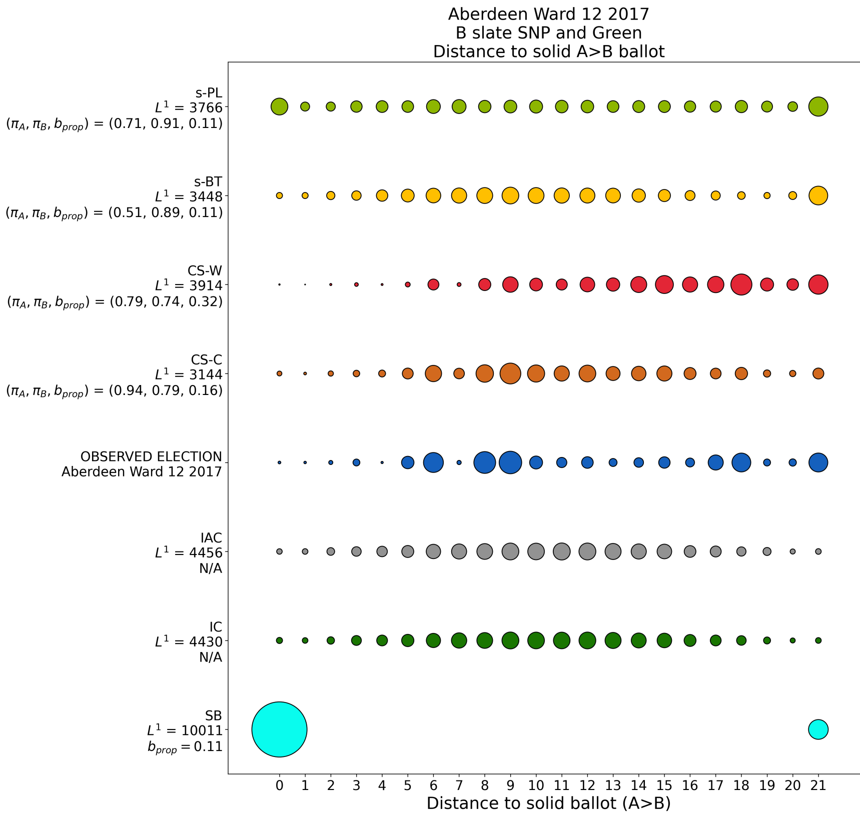

Fig. 10. Bubble plots showing the distribution of swap distances from our generative models, classical models, and a real election to A-over-B ballots. Both the generative models and solid-bloc election are optimized via a grid search to choose a value for π_B that minimizes L^1 to the real Aberdeen Ward 12 2017 election.

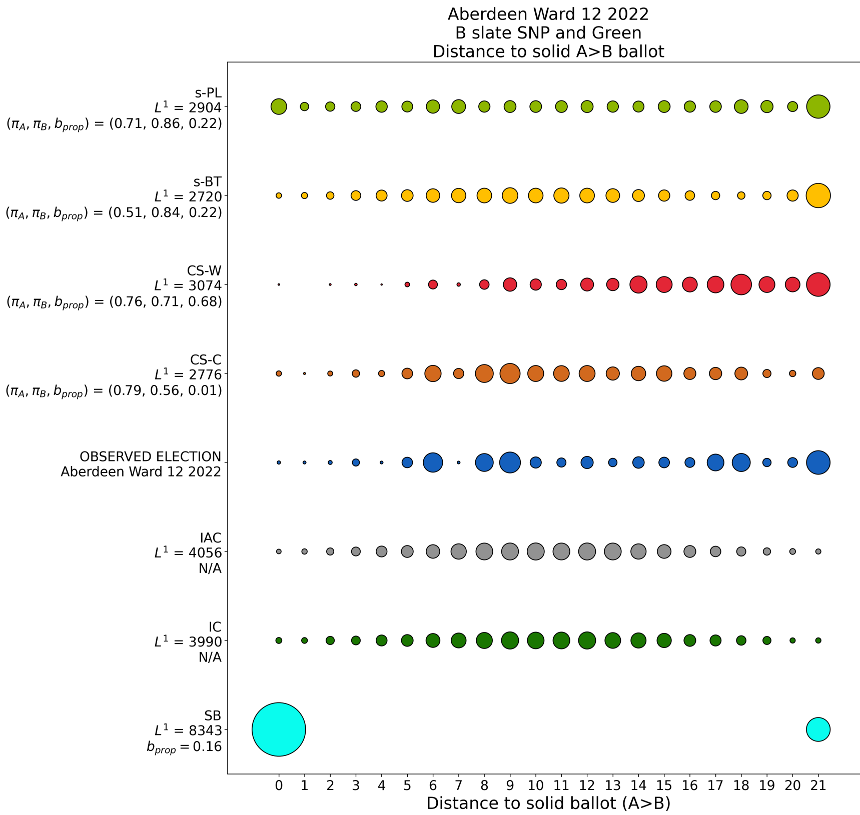

Fig. 11. Bubble plots showing the distribution of swap distances from our generative models, classical models, and a real election to A-over-B ballots. Both the generative models and solid-bloc election are optimized via a grid search to choose a value for π_B that minimizes L^1 to the real Aberdeen Ward 12 2022 election.

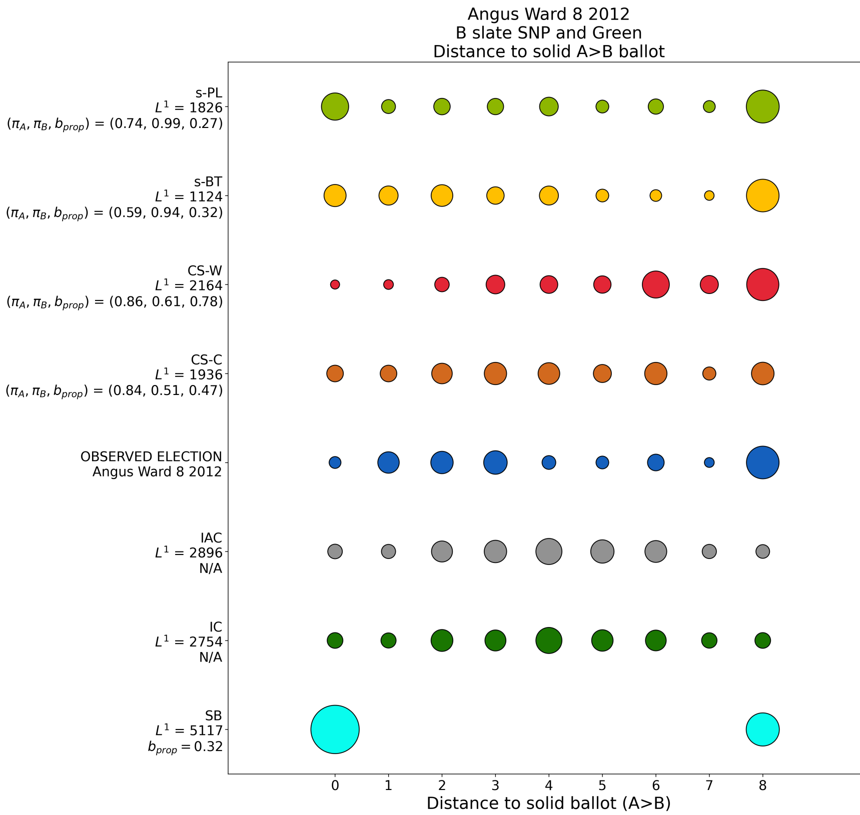

Fig. 12. Bubble plots showing the distribution of swap distances from our generative models, classical models, and a real election to A-over-B ballots. Both the generative models and solid-bloc election are optimized via a grid search to choose a value for π_B that minimizes L^1 to the real Angus Ward 8 2012 election.

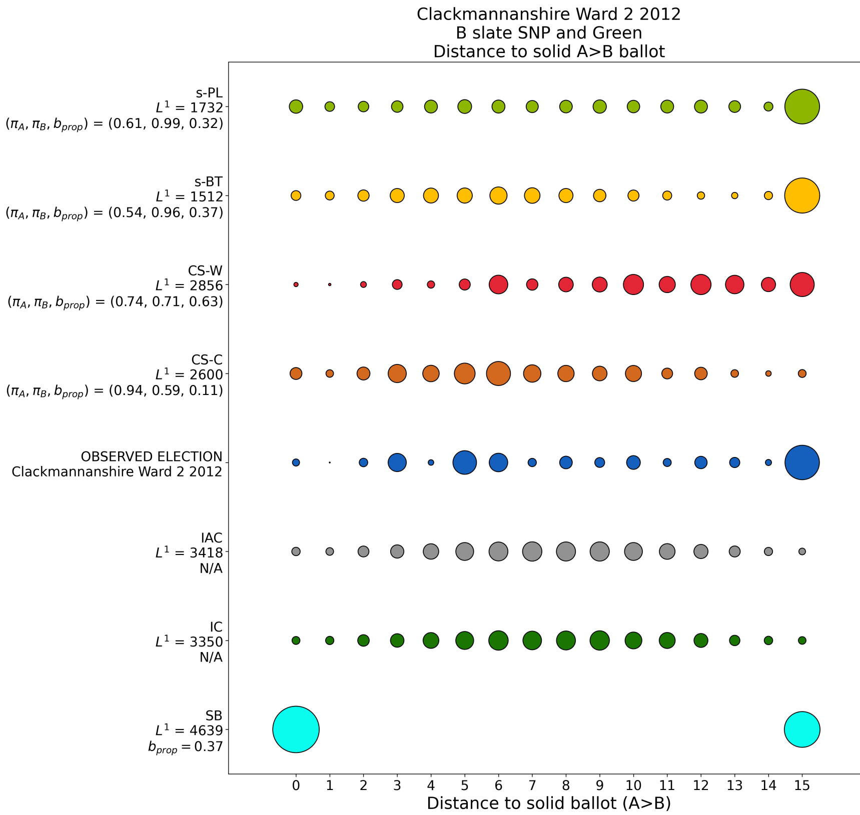

Fig. 13. Bubble plots showing the distribution of swap distances from our generative models, classical models, and a real election to A-over-B ballots. Both the generative models and solid-bloc election are optimized via a grid search to choose a value for π_B that minimizes L^1 to the real Clackmannanshire Ward 2 2012 election.

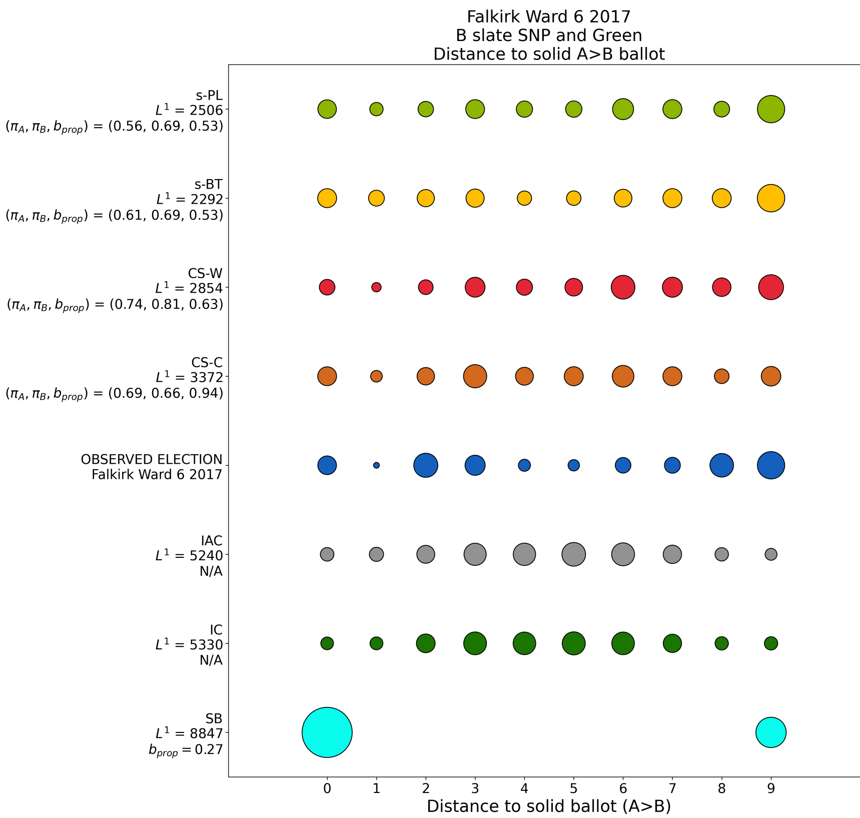

Fig. 14. Bubble plots showing the distribution of swap distances from our generative models, classical models, and a real election to A-over-B ballots. Both the generative models and solid-bloc election are optimized via a grid search to choose a value for π_B that minimizes L^1 to the real Falkirk Ward 6 2017 election.

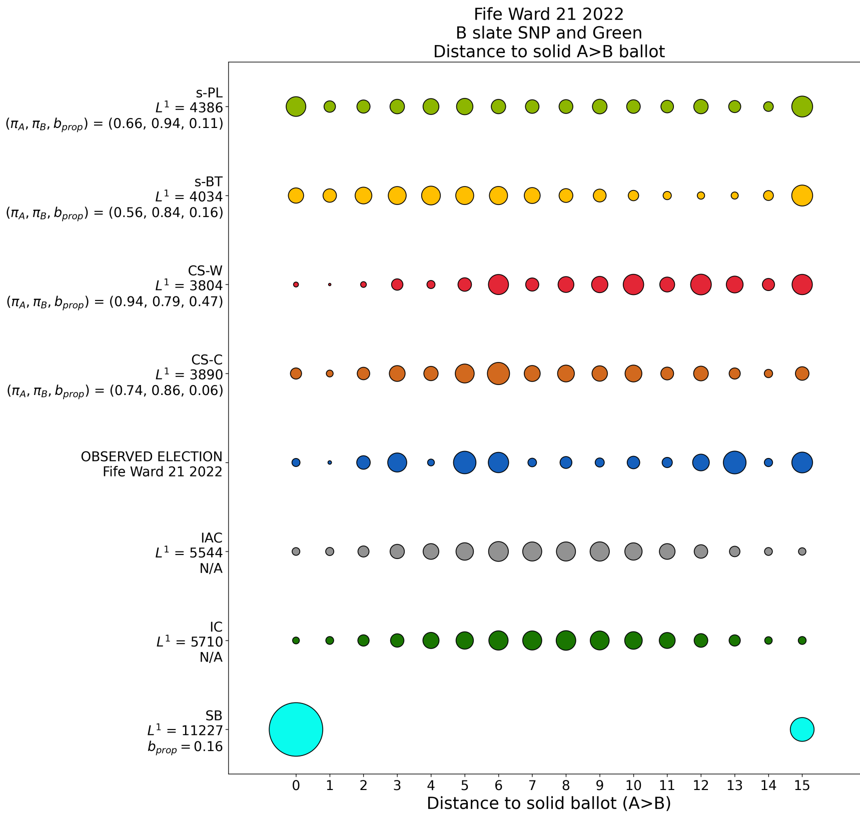

Fig. 15. Bubble plots showing the distribution of swap distances from our generative models, classical models, and a real election to A-over-B ballots. Both the generative models and solid-bloc election are optimized via a grid search to choose a value for π_B that minimizes L^1 to the real Fife Ward 21 2022 election.

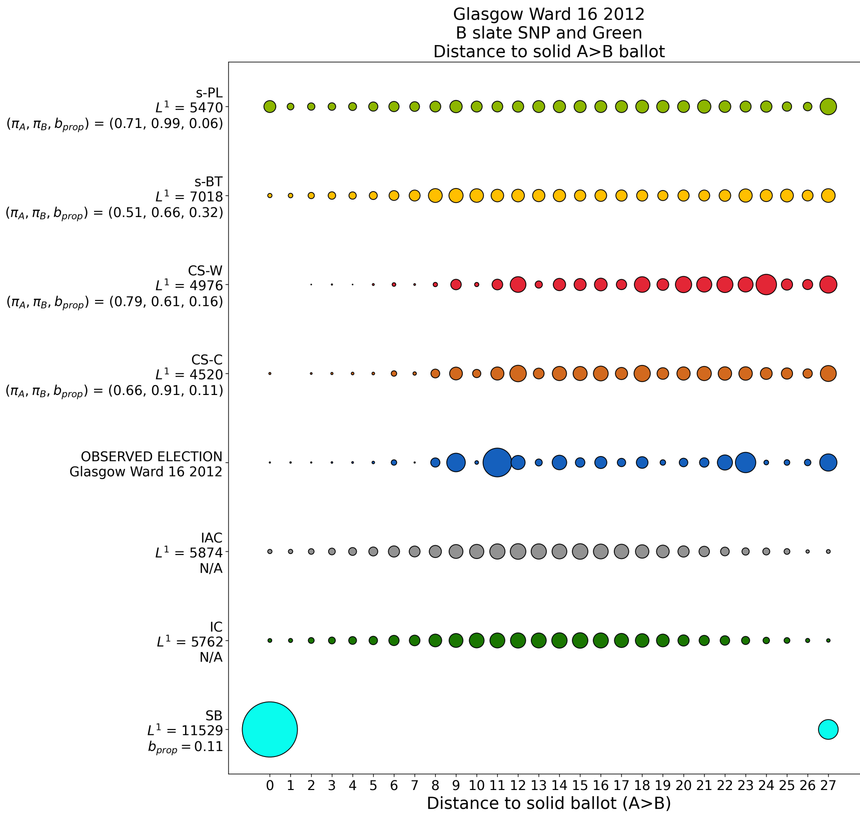

Fig. 16. Bubble plots showing the distribution of swap distances from our generative models, classical models, and a real election to A-over-B ballots. Both the generative models and solid-bloc election are optimized via a grid search to choose a value for π_B that minimizes L^1 to the real Glasgow Ward 16 2012 election.

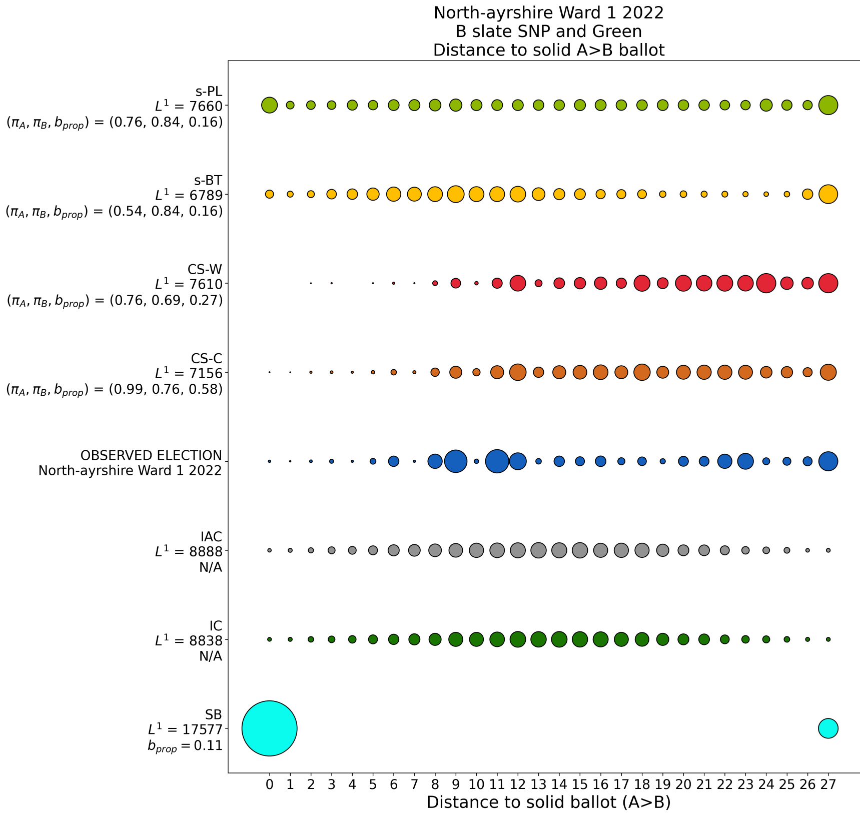

Fig. 17. Bubble plots showing the distribution of swap distances from our generative models, classical models, and a real election to A-over-B ballots. Both the generative models and solid-bloc election are optimized via a grid search to choose a value for π_B that minimizes L^1 to the real North Ayrshire Ward 1 2022 election.

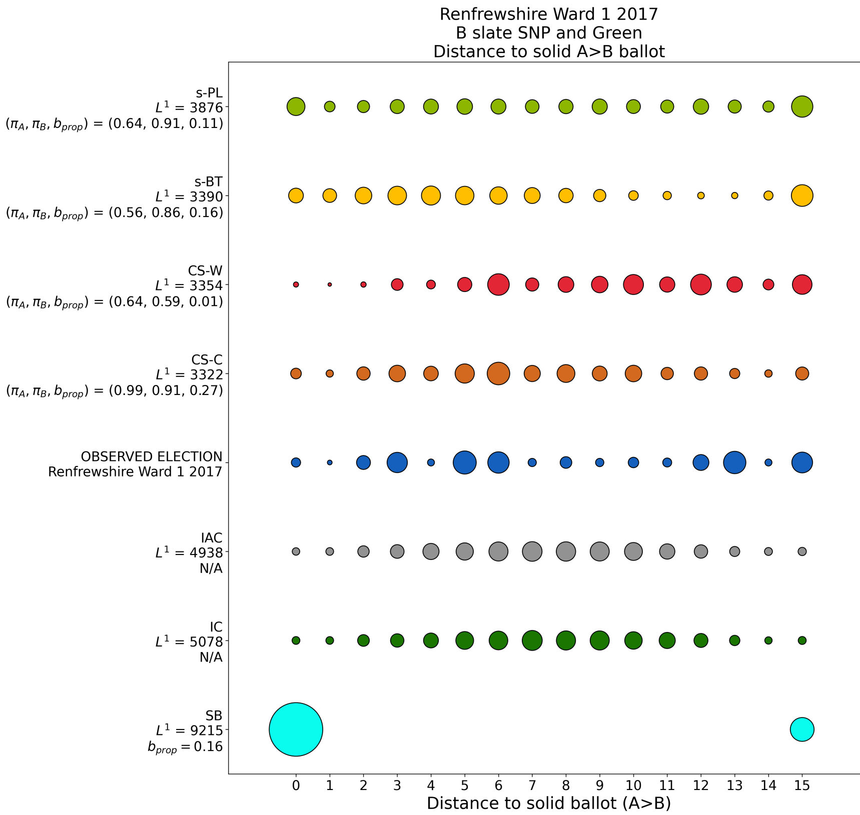

Fig. 18. Bubble plots showing the distribution of swap distances from our generative models, classical models, and a real election to A-over-B ballots. Both the generative models and solid-bloc election are optimized via a grid search to choose a value for π_B that minimizes L^1 to the real Renfrewshire Ward 1 2017 election.