

Agreement, Diversity, and Polarization Indices for Approval Elections

Piotr Faliszewski Jitka Mertlová Krzysztof Sornat Stanisław Szufa Tomasz Wąs

Abstract: An index is a function that given an election outputs a value between 0 and 1, indicating the extent to which this election has a particular feature. We seek indices that capture agreement, diversity, and polarization among voters in approval elections, and that are normalized with respect to saturation. By the latter we mean that if two elections differ by the fraction of candidates approved by an average voter, but otherwise are of similar nature, then they should have similar index values. We propose several indices, analyze their properties, and use them to (a) derive a new map of approval elections, and (b) show similarities and differences between various real-life elections from Pabulib, Preflib and other sources.

Code: <https://github.com/Project-PRAGMA/Approval-DAP-IJCAI-2026>

1 Introduction

Our main goal is to design three functions that given an approval election—i.e., an election where voters give yes/no answers regarding their support for particular candidates—indicate the extent to which the votes are in agreement, are diverse, or are polarized. Following Hashemi and Endriss [2014], we refer to functions that quantify election features (by giving values between 0 and 1) as *election indices*.

Such indices would be valuable because they would allow us to compare approval elections coming from different sources. For example, it is not uncommon to use participatory budgeting (PB) elections from Pabulib [Faliszewski et al., 2023a] outside of the PB context and it is important to validate if this data remains relevant.¹ Further, as recently shown by Faliszewski et al. [2025b], for ordinal elections (where the voters rank the candidates) having good agreement, diversity, and polarization indices leads to appealing maps of elections (i.e., visualizations of election datasets where each election is a dot and distances between these dots give an idea of their elections’ similarity). While a map of approval elections exists [Szufa et al., 2022], it seems to be less useful than those for ordinal ones (see the argument below).

One characteristic issue when dealing with approval elections is that they may largely differ in terms of their *saturation*, i.e., the average number of candidates approved by a voter, while being very similar in nature. As an extreme example, consider one election where all voters approve the same 10% of candidates, and a second one, where all voters approve the same 50% of candidates. Both of these elections represent perfect agreement (and complete lack of diversity and polarization) but according to the map of Szufa et al. [2022], they are quite far from each other. Indeed, that map has two main dimensions, election saturation and the position on the agreement-disagreement spectrum. We seek saturation-independent agreement, diversity, and polarization indices, so that a map based on them would be capable of discovering more nuanced relations between data.

Our general strategy is to focus on designing a saturation-independent agreement index. Given such an index, we derive a diversity one by clustering the voters into as like-minded groups as possible and claiming

Authors’ Information: Piotr Faliszewski , faliszew@agh.edu.pl, AGH University of Kraków, Poland; Jitka Mertlová , mertlova.jita@gmail.com, Czech Technical University in Prague, Czech Republic; Krzysztof Sornat , sornat@agh.edu.pl, AGH University of Kraków, Poland; Stanisław Szufa , s.szufa@gmail.com, University of Geneva, Switzerland; Tomasz Wąs , tomasz.was@cs.ox.ac.uk, University of Oxford, United Kingdom.

¹E.g., for year 2025 Google Scholar points to 18 papers that cite the Pabulib database and perform numerical experiments. Among these, only 12 are indeed focused on participatory budgeting.

that the election is diverse if the agreement within these groups is low. For polarization—which we define as the presence of two groups with conflicting preferences—we take the difference between the overall agreement and the agreements in two voter groups after clustering (the overall agreement in a polarized election would be low, but it would be high in the like-minded groups). This approach was proposed by Faliszewski et al. [2023b] in the ordinal case, for the special case of measuring agreement using the Kemeny rule Kemeny [1959]. We extend their idea.

Altogether, our contributions are as follows. First, we propose a number of agreement indices and analyze them theoretically and experimentally, paying careful attention to saturation independence. In particular, we verify if their results are intuitive on a number of selected election instances. Then, using these indices, we derive and test diversity and polarization ones (using both the general scheme and some other ideas). Finally, we identify three indices—capturing agreement, diversity, and polarization—that work best together in a certain measurable way, and use them to form a map of approval elections. Our map contains both synthetic and real-life elections, e.g., coming from Preflib [Mattei and Walsh, 2013] and Pabulib [Faliszewski et al., 2023a], but also from other sources, including some new ones.

We discuss related work throughout the paper, in relevant places. Omitted proofs and discussions are in the appendix.

2 Preliminaries

For an integer t , we write $[t]$ to denote $\{1, \dots, t\}$. For a vector $x \in \mathbb{R}^t$, we write $x[i]$ to denote its i -th coordinate, and $\bar{x} = \frac{1}{t}(x[1] + \dots + x[t])$ to denote its mean. For $p \in [0, 1]$, $x \sim \text{Bernoulli}(p)$ indicates that x gets value 1 with probability p and value 0 with probability $1 - p$.

Elections. An (*approval*) *election* is a pair $E = (C, V)$, where $C = \{c_1, \dots, c_m\}$ is a set of *candidates* and $V = (v_1, \dots, v_n)$ is a collection of *voters*. Each voter has an *approval ballot* (or, a *vote*) where he or she indicates which candidates he or she approves (and he or she disapproves the other ones). Formally, we represent a vote as a binary vector, whose entries correspond to the candidates and have value 1 if a given candidate is approved and 0 otherwise. To simplify notation, we write v_i to refer both to the respective voter and his or her vote; the exact meaning will always be clear from the context. Additionally, for a voter v_i we write $A(v_i)$ to denote the set of candidates that this voter approves, and for a candidate c_j we write $A(c_j)$ to mean the set of voters approving c_j . Given an election $E = (C, V)$, by E^{-1} we mean an election obtained from E by reversing all entries of all votes (i.e., for each voter v_i and candidate c_j , in E^{-1} , for c_j this voter reports $1 - v_i[j]$). We refer to E^{-1} as a *reverse* of E .

Pearson Correlation Coefficient. Given two t -dimensional vectors, $x, y \in \mathbb{R}^t$, their Pearson correlation coefficient (PCC) is:

$$\text{pcc}(x, y) = \frac{\sum_{i=1}^t (x[i] - \bar{x})(y[i] - \bar{y})}{\sqrt{\sum_{i=1}^t (x[i] - \bar{x})^2} \sqrt{\sum_{i=1}^t (y[i] - \bar{y})^2}}. \quad (1)$$

$\text{pcc}(x, y)$ takes values between 1 and -1 , where 1 means perfect correlation (positive linear dependence between the entries of x and y), -1 means perfect anti-correlation (negative linear dependence between the entries of x and y), and 0 means a complete lack of linear correlation. In rare cases in which in x or y contains only 1 or only 0, the denominator would be equal to 0, so we fix the value of $\text{pcc}(x, y)$ as 1.

(Dis)similarity Measures Between Votes. Given two votes, u and v , over m candidates, their Hamming distance is

$$\text{ham}(u, v) = \sum_{i \in [m]} |u[i] - v[i]|$$

and their Jaccard distances is

$$\text{jacc}(u, v) = 1 - \frac{\sum_{i \in [m]} \min(u[i], v[i])}{\sum_{i \in [m]} \max(u[i], v[i])}.$$

Hamming distance takes values between 0 and m , whereas Jaccard takes values between 0 and 1. We also sometimes measure similarity between two votes using Pearson correlation coefficient, which in the case of binary vectors is known as the phi coefficient Yule [1912], or Matthews coefficient (MCC) Baldi et al. [2000].

There are many other (dis)similarity measures between binary vectors (see, e.g., the discussion of [Duan et al. \[2014\]](#)) but we restrict our attention to these three as they are common in voting literature.

3 Election Data

We test our election indices on three kinds of data: Special, fixed elections of particular, well-understood structure, synthetically generated random elections, and elections obtained from various real-life data (ranging from actual approval elections to those derived from other types of preferences). In the discussion below we consider elections with candidate set $C = \{c_1, \dots, c_m\}$ and voter collection $V = (v_1, \dots, v_n)$.

Special Elections. Our special elections are:

p -Identity (p -ID). In a p -Identity election, all voters approve the same set of $\lfloor pm \rfloor$ candidates (e.g., $c_1, \dots, c_{\lfloor pm \rfloor}$) and disapprove the remaining ones.

k -Party. In a k -Party election there are k disjoint, equal-sized groups of candidates (± 1 , depending on divisibility) and k disjoint, equal-sized groups of voters (± 1 , depending on divisibility). For each $i \in [k]$, the voters in the i -th group approve exactly the candidates in the i -th group. For $x, y \in [0, 1]$, by (x, y) -2-Party we mean a 2-Party election where the first group includes an x fraction of candidates and y fraction of voters, and the second group includes the remaining ones.

For the next three elections we require $n = m$:

Diagonal. This is an m -Party election.

Triangle. For each $i \in [n]$, the i -th voter approves the first i candidates, i.e., $c_1, \dots, c_i$.

Cyclic. The first voter approves the first candidate and last $\lfloor m/2 \rfloor - 1$ candidates. For each $i \in \{2, \dots, n\}$, we obtain the i -th vote by taking the $(i - 1)$ -st one and shifting it cyclically by one position to the right, i.e., replacing each candidate c_j by c_{j+1} , with c_m mapped to c_1 .

Synthetic Elections. Next let us describe our models of generating random elections. If we only describe a process of sampling a single vote, it means that all the votes in the election are generated this way, independently from each other:

p -Impartial Culture (p -IC). For $p \in [0, 1]$, we sample vote v so that for each $i \in [m]$, we have $v[i] \sim \text{Bernoulli}(p)$.

$(p_1, \dots, p_m)$ -Independent Approval Model (IAM). For $(p_1, \dots, p_m) \in [0, 1]^m$, we sample vote v so that for each $i \in [m]$, $v[i] \sim \text{Bernoulli}(p_i)$. IAMs were discussed, e.g., by [Faliszewski et al. \[2025a\]](#), [Lackner and Maly \[2025\]](#), and [Xia \[2025\]](#).

(p, ϕ) -Resampling. Given $p, \phi \in [0, 1]$ and a fixed central vote u with $|A(u)| = \lfloor pm \rfloor$, we sample vote v as follows: For each $i \in [m]$, with probability $1 - \phi$ we let $v[i] = u[i]$ and with probability ϕ we let $v[i] \sim \text{Bernoulli}(p)$. This model is due to [Szufa et al. \[2022\]](#). Some authors, such as [Faliszewski et al. \[2025a\]](#), allow arbitrary central votes, but we use the original variant as it gives pm approved candidates in expectation.

Euclidean. Here, we assume that each voter and candidate is associated with a point in a d -dimensional Euclidean space, and each voter v_i approves a candidate c if the distance between their corresponding points is smaller than some threshold value r_i (which can be different for each voter). Such a model was studied, e.g., by [Godziszewski et al. \[2021\]](#). See [Section D](#) for details.

We also consider a few special distributions, to generate elections with particular features:

p -ID/IC. The first half of the voters form a p -ID election and the second half is generated using the p -IC model.

Lin-IC. In this model, v_1 is generated using 1-IC, v_2 using $(1 - 1/n)$ -IC, v_3 using $(1 - 2/n)$ -IC, and so on, until v_n , generated using $1/n$ -IC.

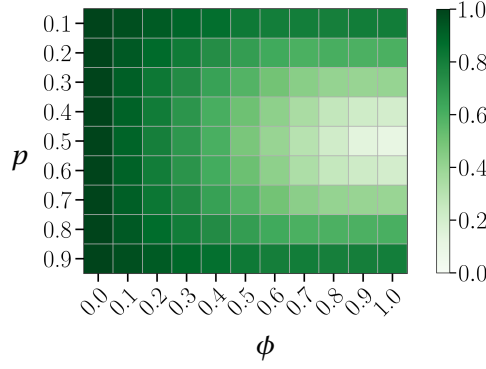


Figure 1: Results of the resampling experiment for Approval Agreement index. Rows correspond to fixed p values and columns to fixed ϕ values.

$N(\mathcal{D}, \phi)$. $\mathcal{D}$ is a distribution over elections. We first generate an election using $\mathcal{D}$ and then we replace each vote v with a new one as follows: We let p be the fraction of candidates approved in v and we generate a vote using (p, ϕ) -resampling model with central vote equal to v . In other words, $N(\mathcal{D}, \phi)$ introduces noise on top of elections generated using $\mathcal{D}$ (N stands for *noisy*).

Real-Life Data. See [Section 7](#) and [Section D](#).

4 Indices and Initial Experiments

By an election index, we mean a function f that given an election E outputs a value between 0 and 1, quantifying the extent to which E has a feature captured by f . In the following sections we define agreement, diversity, and polarization indices. For now, we discuss indices on a high level and describe our approach to evaluating them.

Normalization. Given an election $E = (C, V)$, we consider its average vote length (avl; the average number of candidates approved by a voter), reverse avl (the average number of candidates disapproved by a voter), and its saturation (satr; the fraction of candidates that an average voter approves):

$$\text{avl}(E) = \frac{1}{|V|} \sum_{v \in V} |A(v)|, \quad \text{rev-avl}(E) = \text{avl}(E^{-1}), \quad \text{satr}(E) = \text{avl}(E)/|C|.$$

While saturation is a natural election index by itself, for the other indices we are mostly interested in normalizing them in a way that would be independent of saturation. In other words, if two elections are similar in nature except that their saturations differ (as would be the case, e.g., for 0.5-IC and 0.25-IC elections), we would prefer an election index to give them similar values. Naturally, this is not a precise mathematical statement, but rather an intuitive desideratum. Yet, in some cases it is easy to argue that a particular index is not saturation-independent, and we often view it as an argument against this index.

Resampling Experiment. For each of our indices, we perform the following *resampling experiment*: For each $p \in \{0.1, \dots, 0.9\}$ and each $\phi \in \{0, 0.1, \dots, 1\}$, we generate 10 (p, ϕ) -resampling elections with 60 candidates and 60 voters, and compute for them the average value of the index in question. We arrange the results as a matrix, where each row corresponds to a fixed value of p and each column corresponds to a fixed value of ϕ . Color brightness represents the average value of the index. We present such matrices in the first row of [Table 1](#), and we provide an enlarged example (including axis labels and the color-bar scale) in [Figure 1](#). The important feature of these plots is that each column corresponds to a fixed value of ϕ , while saturation (i.e., p) varies across rows. Hence, if the columns of the matrix are not (approximately) uniformly colored, then we claim that a given index is not saturation-independent (as, arguably, resampling elections with different values of p but the same value of ϕ are of the same nature). For such an index, we say that it *fails the resampling test of saturation independence*.

		av-agr	cntr-agr	pair-agr	pcc-agr	jacc-agr	pcc ⁺ -agr	cntr-div	pcc-div	out-div	cntr-pol	pcc-pol	pair-pol
resampling experiment:													
saturation independence:		✗	✗	✓	✓	✗	✓	✗	✓	✓	—	—	—
1.	1/3-ID		1.00 ±0.00	1.00 ±0.00	1.00 ±0.00	1.00 ±0.00	1.00 ±0.00	0.00 ±0.00	0.00 ±0.00	0.00 ±0.00	0.00 ±0.00	0.00 ±0.00	0.00 ±0.00
2.	2-Party		0.00 ±0.00	0.00 ±0.00	0.00 ±0.00	0.00 ±0.00	0.50 ±0.00	0.20 ±0.00	0.25 ±0.02	0.10 ±0.00	1.00 ±0.00	1.00 ±0.00	1.00 ±0.00
3.	N(2-Party, 0.6)		0.09 ±0.01	0.08 ±0.01	0.01 ±0.00	0.02 ±0.00	0.35 ±0.01	0.66 ±0.02	0.82 ±0.01	0.29 ±0.00	0.25 ±0.10	0.17 ±0.01	0.24 ±0.01
4.	3-Party		0.33 ±0.00	0.00 ±0.00	0.00 ±0.00	0.00 ±0.00	0.33 ±0.00	0.33 ±0.00	0.31 ±0.01	0.13 ±0.00	0.33 ±0.00	0.50 ±0.00	0.63 ±0.00
5.	4-Party		0.50 ±0.00	0.00 ±0.00	0.00 ±0.00	0.00 ±0.00	0.25 ±0.00	0.47 ±0.04	0.40 ±0.00	0.15 ±0.00	0.25 ±0.00	0.33 ±0.00	0.43 ±0.00
6.	(1/3, 1/3)-2-Party		0.33 ±0.00	0.24 ±0.00	0.10 ±0.00	0.11 ±0.00	0.56 ±0.00	0.15 ±0.00	0.18 ±0.00	0.09 ±0.00	0.76 ±0.00	0.89 ±0.00	0.99 ±0.00
7.	Cyclic		0.00 ±0.00	0.00 ±0.00	0.00 ±0.00	0.00 ±0.00	0.39 ±0.00	0.46 ±0.00	0.57 ±0.01	0.22 ±0.00	0.50 ±0.00	0.33 ±0.00	0.58 ±0.00
8.	Diagonal		0.97 ±0.00	0.00 ±0.00	0.00 ±0.00	0.00 ±0.00	0.02 ±0.00	0.97 ±0.00	0.97 ±0.00	0.63 ±0.01	0.02 ±0.00	0.02 ±0.00	0.01 ±0.00
9.	Triangle		0.50 ±0.00	0.49 ±0.00	0.33 ±0.00	0.50 ±0.00	0.51 ±0.00	0.41 ±0.00	0.32 ±0.00	0.18 ±0.00	0.01 ±0.00	0.14 ±0.00	0.47 ±0.00
10.	N(Tri, 0.6)		0.21 ±0.01	0.20 ±0.02	0.06 ±0.00	0.15 ±0.02	0.33 ±0.01	0.89 ±0.01	0.77 ±0.03	0.26 ±0.00	0.00 ±0.02	0.05 ±0.00	0.39 ±0.01
11.	1/2-ID/IC		0.50 ±0.01	0.49 ±0.01	0.26 ±0.01	0.25 ±0.01	0.51 ±0.00	0.41 ±0.01	0.50 ±0.00	0.19 ±0.00	0.07 ±0.01	0.26 ±0.02	0.45 ±0.01
12.	1/2-IC		0.10 ±0.01	0.10 ±0.01	0.02 ±0.00	0.02 ±0.00	0.34 ±0.01	0.81 ±0.01	0.91 ±0.00	0.29 ±0.00	0.07 ±0.01	0.05 ±0.00	0.18 ±0.00
13.	1/4-IC		0.49 ±0.01	0.00 ±0.00	0.02 ±0.00	0.02 ±0.00	0.16 ±0.00	0.97 ±0.00	0.90 ±0.00	0.31 ±0.00	0.01 ±0.01	0.05 ±0.00	0.16 ±0.00
14.	Lin-IC		0.08 ±0.01	0.06 ±0.01	0.01 ±0.00	0.09 ±0.02	0.31 ±0.01	0.96 ±0.01	0.83 ±0.03	0.27 ±0.00	0.00 ±0.01	0.04 ±0.00	0.37 ±0.01

Table 1: Values of our indices for elections from various distributions. The elections have 60 candidates and 60 voters. Index values are averaged over 10 samples and the standard deviation is reported in smaller font. Similar table with more elections can be found in [Section D](#).

Index Values on Example Elections. In [Table 1](#) we also give values of our indices on a number of characteristic (distributions over) elections. Each row gives index values for a given family of elections. All elections have 60 candidates and 60 voters and the results for each family are averaged over sampling 10 elections (we report standard deviation in smaller font). For each family we give its name and an icon presenting an example election; in this icon, columns are the candidates, rows are the voters, and darker spots indicate that a given voter approves a given candidate. We discuss this table throughout the following sections.

5 Agreement Indices

An agreement index should identify whether the voters have similar views, i.e., tend to approve and disapprove the same candidates. We consider two types of such indices: Global ones look at the extent to which all voters agree with each other, whereas local ones also detect large groups of voters that are internally in agreement, but that may disagree with each other. We propose six agreement indices that we describe below. All our indices give value 1 exactly for p -ID elections—which, indeed, represent perfect agreements—and for most of them we also characterize when they give value 0.

5.1 Approval Agreement

Our first index, Approval Agreement, adapts an idea from the world of ordinal elections (where the voters rank the candidates). There, for every pair of candidates a and b we compute the absolute difference between the fractions of voters that rank them one way or the other, and output the average of these values [[Alcalde-Unzu and Vorsatz, 2013](#); [Hashemi and Endriss, 2014](#); [Can et al., 2015](#); [Faliszewski et al., 2023b](#)]. In the approval

setting, we compare the numbers of voters approving and disapproving each given candidate.

Definition 5.1. *Given an election $E = (C, V)$, its Approval Agreement is*

$$\text{av-agr}(E) = \frac{1}{|C|} \sum_{c \in C} \left| \frac{|A(c)|}{|V|} - \left(1 - \frac{|A(c)|}{|V|} \right) \right| = \frac{1}{|C|} \sum_{c \in C} \left| 1 - 2 \frac{|A(c)|}{|V|} \right|.$$

Unfortunately, this index has several drawbacks. Foremost, as we see in the first row of [Table 1](#), it fails the resampling test of saturation independence (and, indeed, we see that it gives different values for $1/2$ -IC and $1/4$ -IC). This is also visible in the following proposition, where we show that it assumes value 0 if and only if each candidate is approved by exactly half of the voters and, hence, it is not possible to assume this value for saturation different from 0.5.

Proposition 5.1. *For an election E , $\text{av-agr}(E) = 1$ exactly if E is an identity election, and $\text{av-agr}(E) = 0$ exactly if each candidate is approved by exactly half of the voters.*

The index also behaves poorly on Party elections. As we see in [Table 1](#), for 2-Party its value is 0, but it increases as we consider 3-Party and 4-Party, to become close to 1 for Diagonal. Indeed, the index treats the fact that many voters disapprove the same candidates as a sign of agreement among them (while in some cases—such as elections within parliaments—this indeed might be a reasonable view, in others—such as in participatory budgeting—it is ungrounded).

5.2 Central Agreement(s)

Our second index, Central Agreement, is designed to resemble the Kemeny voting rule for ordinal elections [[Kemeny, 1959](#)]. In that setting, a Kemeny ranking is a ranking that minimizes the sum of swap distances to all the votes in a given election. The larger is this value, the bigger is the disagreement among the voters. We implement the same idea for approval elections using the Hamming distance (while computing a Kemeny ranking is well-known to be intractable [[Bartholdi III et al., 1989](#); [Hemaspaandra et al., 2005](#)], in our case the analogous task is very simple).

Let $E = (C, V)$ be an election. A vote u over candidate set C is central for E if for each candidate $c \in C$, u agrees regarding c with at least half of the voters. So, if exactly half of the voters approve some candidate and half disapprove it, then there is more than one central vote for the given election. We write $\text{cen}(E)$ to denote the set of all central votes for E .

Definition 5.2. *Let $E = (C, V)$ be an election and let u be an arbitrary member of $\text{cen}(E)$. We define*

$$\text{cntr-agr}(E) = 1 - \frac{\sum_{v_i \in V} \text{ham}(v_i, u)}{|V| \cdot \min(\text{avl}(E), \text{rev-avl}(E))}.$$

The denominator in the definition of $\text{cntr-agr}(E)$ is chosen to be (nearly) equal to the largest possible value of the numerator for elections of the same size and saturation as E .

Proposition 5.2. *For an election E , $\text{cntr-agr}(E) = 1$ exactly if E is an identity election, and $\text{cntr-agr}(E) = 0$ exactly if either $\text{cen}(E)$ contains a vote approving all candidates or $\text{cen}(E)$ contains a vote disapproving all candidates.*

Central Agreement has two major drawbacks. First, as indicated by [Theorem 5.2](#), it is quite non-discriminative and often assumes value 0 (e.g., on the majority of Pabulib elections). Second, it fails the resampling test of saturation independence (see [Table 1](#)). Interestingly, Central Agreement and Approval Agreement differ mostly by their normalization (replacing the denominator with m in [Theorem 5.3](#) below, we would get Approval Agreement).

Proposition 5.3. *Let $E = (C, V)$, then*

$$\text{cntr-agr}(E) = 1 - \sum_{c \in C} \frac{1 - \left| 1 - \frac{2|A(c)|}{|V|} \right|}{2 \cdot \min(\text{avl}(E), \text{rev-avl}(E))}.$$

We do not consider central agreement based on the Jaccard distance as it would be intractable [[Chierichetti et al., 2010](#)] and we find sufficiently many other agreement indices.

5.3 Pairwise Agreements

The final four indices are based on computing expected (dis)similarity between two randomly selected votes.

Definition 5.3. For an election $E = (C, V)$, we define Hamming, Jaccard, PCC, and PCC^+ Pairwise Agreement Indices:

$$\begin{aligned}\text{pair-agr}(E) &= 1 - \frac{\sum_{u \in V} \sum_{v \in V} \text{ham}(u, v)}{2|V|^2|C| \cdot \text{satr}(E)(1 - \text{satr}(E))}, \\ \text{jacc-agr}(E) &= \frac{1}{|V|^2} \sum_{u \in V} \sum_{v \in V} (1 - \text{jacc}(u, v)), \\ \text{pcc-agr}(E) &= \frac{1}{|V|^2} \sum_{u \in V} \sum_{v \in V} \text{pcc}(u, v), \\ \text{pcc}^+ \text{-agr}(E) &= \frac{1}{|V|^2} \sum_{u \in V} \sum_{v \in V} \max(0, \text{pcc}(u, v)).\end{aligned}$$

The normalization for Hamming Pairwise Agreement follows that for Central Agreement: the denominator is equal to the maximum possible value of the numerator among elections with the same size and saturation (assuming that it is possible to ensure that every candidate has the same approval score). The normalizations for Jaccard and PCC^+ use the fact that we average over values that are between 0 and 1. For PCC Agreement, we need an additional argument.

Proposition 5.4. For each election E , $\text{pcc-agr}(E) \in [0, 1]$.

At first sight, computing each of these indices requires $O(|C||V|^2)$ time: Computing similarity between a pair of votes requires $O(|C|)$ time and there are $O(|V|^2)$ pairs to consider. Yet, by changing the order of summation, we can compute Hamming Pairwise Agreement in $O(|C||V|)$ time.

Theorem 5.5. For an election $E = (C, V)$, we have that

$$\text{pair-agr}(E) = 1 - \frac{\sum_{c \in C} |A(c)| \cdot (|V| - |A(c)|)}{|V|^2|C| \cdot \text{satr}(E)(1 - \text{satr}(E))}.$$

Next we characterize when Hamming and Jaccard Pairwise indices assume values 1 and 0. In case of PCC-based indices we only provide some examples and observations.

Theorem 5.6. Each of Pairwise, Jaccard, PCC, and PCC^+ Agreement Indices assumes value 1 exactly for identity elections. Hamming Pairwise Agreement has value 0 exactly if all candidates have equal approval score, and Jaccard Pairwise Agreement has value 0 exactly if all the voters approve disjoint sets of candidates.

Proposition 5.7. For each k -Party election $E = (C, V)$ such that both $|C|$ and $|V|$ are divisible by k , we have $\text{pcc-agr}(E) = 0$ and $\text{pcc}^+ \text{-agr}(E) = 1/k$.

Remark 5.1. Let $p \in [0, 1]$ be a number and let u and v be two votes such that a p^2 fraction of candidates is approved in both votes, a $(1 - p)^2$ fraction is disapproved in both, a $p(1 - p)$ fraction is only approved in u , and a $p(1 - p)$ fraction is only approved in v . Then $\text{pcc}(u, v) = 0$. As votes in p -IC elections are close to satisfying these conditions, we expect PCC and PCC^+ Agreements to be close to 0 for IC elections.

We note that Hamming Pairwise Agreement and PCC Agreement tend to give very similar results for the examples in Table 1. In fact, if all voters approve the same number of candidates, then the two indices are equal (but this is not the case when vote lengths differ; see the Triangle election).

Proposition 5.8. Let $E = (C, V)$ be an election such that $0 < |A(u)| = |A(v)| < |C|$ for all $u, v \in V$. Then we have that $\text{pair-agr}(E) = \text{pcc-agr}(E)$.

In the resampling experiment both Hamming and PCC Agreements give results independent of saturation. Overall, we view both of them as good global agreement indices.

Jaccard and PCC^+ Pairwise Agreement indices capture local agreement, by detecting large agreeing groups, even if they disagree with other such groups (indeed, in PCC^+ we only sum positive PCC values to account for like-minded voters, disregarding negative correlations that could cancel them out). For example, they give high—and nearly identical—values for party elections (the fewer parties, the higher agreement, as there are larger agreeing groups) and detect agreeing groups in $1/2$ -ID/IC, and Lin-IC. Yet, we prefer PCC^+ Agreement as it passes the resampling test of saturation independence and assigns close-to-zero values to IC elections. The fact that Jaccard Pairwise Agreement fails the resampling test of saturation is expected, as it treats approvals and disapprovals asymmetrically. If such asymmetry is natural in a given setting it still might be a valuable index.

6 Diversity and Polarization Indices

Intuitively, an election is diverse if (a) its votes tend to be different from each other, and (b) they cover the space of all possible votes (modulo saturation). For example, we view Diagonal as a maximally diverse election, because for its saturation (one approval per voter) it contains exactly one copy of each possible vote. Yet, we view Cyclic, which is similar in spirit to Diagonal, as much less diverse than $1/2$ -IC; these two types of election have the same saturation, but the votes in the latter are notably more different from each other and, intuitively, cover the preference space more evenly.

On the other hand, we say that an election is polarized if it consists of two groups of voters that agree internally but strongly disagree with each other. Consequently, 2-Party is an example of a maximally polarized election. We acknowledge that elections with three, four, or more internally consistent but disagreeing groups could also be seen as maximally polarized—and seeking indices capturing this intuition would be valuable—but we leave this for future work.

For a broader view of diversity and polarization, see the special issue edited by [Levin et al. \[2021\]](#). In voting literature, see, e.g., the works of [Nehring and Puppe \[2002\]](#), [Hashemi and Endriss \[2014\]](#), [Can et al. \[2015\]](#), [Karpov et al. \[2024\]](#), [Ammann and Puppe \[2025\]](#), and [Faliszewski et al. \[2023b; 2025b; 2026a; 2026b\]](#).

6.1 Clustering-Based Diversity and Polarization

Our first approach to designing diversity indices follows the view of [Faliszewski et al. \[2023b, 2026a\]](#), that an election is diverse if even after clustering its voters into like-minded groups, these groups disagree internally. As opposed to them, we parameterize our definition with an agreement index.

For a given integer k and an election $E = (C, V)$, let $\mathcal{P}_k(V)$ be a set of all complete partitions of voters into k disjoint subsets $V_1, \dots, V_k$, and let a be some agreement index. We aim to find a partition that maximizes average agreement within the subsets, weighted by the size of each cluster, i.e., $\sum_{i=1}^k \frac{|V_i|}{|V|} a(C, V_i)$. In principle, there are many ways in which the values for different k can be aggregated to give a single index. We follow a simple approach of [Faliszewski et al. \[2025b\]](#) and take the average of values for k from 1 to 5. This results in the following definition.

Definition 6.1. For an election $E = (C, V)$, and agreement index a , we define the a -Diversity index as:

$$a\text{-div}(E) = 1 - \frac{1}{5} \sum_{k=1}^5 \max_{\{V_1, \dots, V_k\} \in \mathcal{P}_k(V)} \sum_{i=1}^k \frac{|V_i|}{|V|} a(C, V_i).$$

Taking Central Agreement as the underlying index, we obtain Central Diversity, denoted cntr-div , and taking PCC Pairwise Agreement we obtain PCC Diversity, denoted pcc-div . We do not consider Hamming Pairwise Agreement for the sake of focus and due to its similarity to pcc-agr . In both cases, obtaining the optimal partition of votes is not computationally feasible, thus we rely on clustering heuristics. We approximate cntr-div with an algorithm based on the k -means clustering: We employ the k -medoids version using Hamming distance, where cluster centers are restricted to observed votes and are iteratively updated to minimize the sum of intra-cluster distances. For pcc-div we use spectral clustering [\[Ng et al., 2001\]](#); our input for spectral clustering is the number k of clusters and the affinity matrix, where for each pair of votes u, v we provide value $1 + \frac{1}{2}\text{pcc}(u, v)$ (value 0 means total dissimilarity, and 1 means the votes are equal).

To quantify polarization, we measure the increase of agreement arising from clustering the election into two groups instead of viewing it as a single group (again, the general idea is due to [Faliszewski et al. \[2023b, 2026a\]](#)).

Definition 6.2. For an election $E = (C, V)$, and agreement index a , we define the a -Polarization index as:

$$a\text{-pol}(E) = \max_{\{V_1, V_2\} \in \mathcal{P}_2(V)} \sum_{i=1}^2 \frac{|V_i|}{|V|} a(C, V_i) - a(E).$$

We consider Central Polarization (cntr-pol) and PCC Polarization (pcc-pol), based on the Central and PCC Pairwise Agreement indices, and we use the same clustering heuristics as in the case of diversity indices.

Additionally, we consider the Pairwise Polarization index, which outputs the standard deviation of Hamming distances between all pairs of votes. The intuition is that in a polarized electorate each two voters either have very similar or close-to-opposite preferences (factor of $2/|C|$ ensures values in $[0, 1]$).

Definition 6.3. For an election $E = (C, V)$, the Pairwise Polarization index is defined as

$$\text{pair-pol}(E) = 2\sqrt{\frac{\sum_{u,v \in V} (\text{ham}(u,v) - \sum_{x,y \in V} \text{ham}(x,y)/|V|^2)^2}{|V|^2|C|^2}}.$$

6.2 Outer Diversity

Recently, [Faliszewski et al. \[2026b\]](#) proposed an alternative diversity index, based on the average distance between the votes in the election and all possible votes. If the average distance is small, the votes are well spread in the whole space, i.e., diverse. It was originally proposed in the ordinal setting and for domains (sets of votes without specified multiplicity) rather than elections, but we adapt it to our setting as follows.

For two collections of votes over the same set of candidates, $U = (u_1, \dots, u_k)$ and $V = (v_1, \dots, v_n)$, by $\text{ham}(U, V)$ we mean the average Hamming distance between n copies of votes in U and k copies of votes in V , matched so as to minimize this distance, i.e.:

$$\text{ham}(U, V) = \frac{1}{kn} \min_{\pi: [kn] \rightarrow [kn]} \sum_{i=1}^{kn} \text{ham}(u_i, v_{\pi(i)}),$$

where π is a bijection and indices from U are taken modulo k and from V modulo n .

For a candidate set C , let $\mathcal{U}_C$ consist of a single copy of each vote from $\{0, 1\}^{|C|}$. Then, in principle, we could measure diversity of an election $E = (C, V)$ based on $\text{ham}(V, \mathcal{U}_C)$, i.e., the average distance between votes in V and all possible votes. However, if the saturation of E is very small or large, the value of $\text{ham}(V, \mathcal{U}_C)$ is always substantial (as most votes in $\mathcal{U}_C$ have moderate saturation). Thus, no election with small/large saturation could be considered diverse according to such a measure. To circumvent this issue, instead of $\mathcal{U}_C$ we use $p\text{-}\mathcal{U}_C$, defined as a collection in which the number of copies of each vote is proportional to the probability of drawing such a vote from $p\text{-IC}$ (so, $1/2\text{-}\mathcal{U}_C = \mathcal{U}_C$).

We would like the index to be maximum and equal to 1 if $V = p\text{-}\mathcal{U}_C$, and minimum and equal to 0 if V is a singleton. From the following lemma we get that for a single vote v with saturation p , it holds that $\text{ham}(\{v\}, p\text{-}\mathcal{U}_C) = 2p(1-p)$.

Lemma 6.1. For candidate set C , $q \in \{0, \frac{1}{|C|}, \dots, 1\}$ and $v \in \{0, 1\}^{|C|}$ with $q|C|$ ones it holds that:

$$\text{ham}(\{v\}, p\text{-}\mathcal{U}_C) = p(1-q) + q(1-p).$$

Definition 6.4. For an election $E = (C, V)$ with $\text{satr}(E) = p$, the Outer Diversity index is defined as

$$\text{out-div}(E) = 1 - \frac{1}{2p(1-p)} \text{ham}(V, p\text{-}\mathcal{U}_C).$$

As discussed, the value 1 obtained for $V = p\text{-}\mathcal{U}_C$, and a single vote v always yields the value of 0. It remains to prove that this is indeed the minimum and maximum value.

Proposition 6.2. For each election E , $\text{out-div}(E) \in [0, 1]$.

6.3 Comparison of the Indices

Let us now analyze the values of the indices given in [Table 1](#). First, we see that Central Diversity fails the resampling test of saturation independence, whereas the other diversity indices pass it. For polarization indices the test is not meaningful, as we expect low polarization for all values of p and ϕ parameters. We also see that clustering-based diversity indices give close-to-1 values for Diagonal and IC elections (including Lin-IC). Outer diversity performs surprisingly poorly on these elections, giving them low values, but we have verified that its values are much closer to 1 for analogous elections with a notably larger number of voters. Still, this dependence on election size seems a major drawback of the index. This is quite disappointing, given how well the outer diversity idea worked for ordinal preference domains [\[Faliszewski et al., 2026b\]](#). Further, all diversity indices give higher values for the noisy variants of respective elections (see the values of 2-Party versus N(2-Party, 0.6) and Triangle versus N(Triangle, 0.6)).

All polarization indices give value 1 for 2-Party and lower values for its noisy variant, as well as for 3- and 4-Party, as expected. Pairwise Polarization still gives nearly-1 for $(1/3, 1/3)$ -2-Party whereas Central and PCC Polarization give lower ones, but both approaches could be justified. Overall, Pairwise Polarization tends to give highest values. For example, for Triangle it gives nearly $1/2$, while the other indices give values much closer to 0. Both approaches could be justified, either by Triangle's similarity to 2-Party, or by saying that in Triangle all voters agree on best-to-worst ranking of the candidates, but have different approval thresholds.

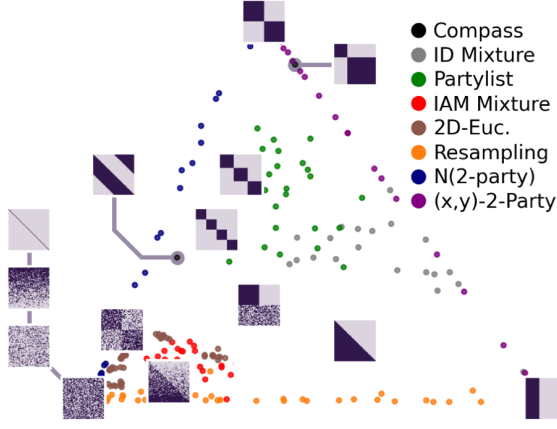


Figure 2: Map of synthetic elections.

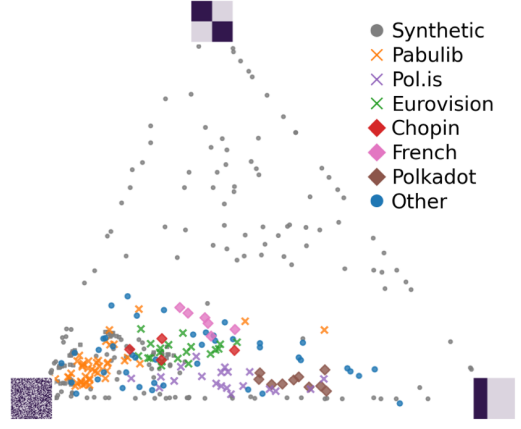


Figure 3: Map of real-life elections.

7 Experiments

To present experimental results for our indices, we use the *map of elections* framework [Szufa et al., 2025]. We construct the map as follows. First, we assemble a collection of elections. Next, we compute pairwise distances between elections using the feature distance, as in the work of Faliszewski et al. [2025b]. Specifically, let $f(E) = (f_1(E), \dots, f_k(E))$ denote the feature vector of election E , where $f_1, \dots, f_k$ are our indices. For two elections E and F , we define their distance as the (standard) Euclidean distance between their feature vectors, $d_f(E, F) = \|f(E) - f(F)\|_2$. Finally, we embed the resulting distance matrix using multidimensional scaling (MDS [Kruskal, 1964]).

For the map, we use the following three indices: pcc-agr, pcc-div, and pcc-pol.² In Figure 2, we show a map based on synthetic elections, while Figure 3 presents a map based on real-life ones (all our datasets will be publicly available). Each point corresponds to a single election. We describe the selected/generated elections in the Section D.

Interestingly, the synthetic map forms a triangle whose vertices correspond to three extreme cultures: ID (maximizing agreement), IC (maximizing diversity), and 2-party (maximizing polarization). The edge between IC and ID is populated by Resampling elections; the edge between 2-party and IC is filled by N(2-Party) elections; and the edge between 2-party and ID is filled by (x, y) -2-Party elections. Both 2D-Euclidean (brown) and IAM (red) elections lie close to IC, reflecting high diversity together with low polarization and agreement. Finally, the partylist (green) elections occupy the upper region of the triangle, indicating high polarization.

Turning to the real-life map (Figure 3), we observe that most elections exhibit very low polarization and, on average, substantially higher diversity than agreement. Compared with other real-world datasets, the Polkadot blockchain data (stakeholders voting over validators) shows relatively high agreement. Both Polkadot and Pol.is (users approving comments on an online deliberation platform) exhibit particularly low polarization. Intuitively, this may reflect a shared “search for ground truth” dynamic in both settings. Notably, the Eurovision (song contest) data looks similar to the Chopin (piano competition) one, which is interesting given that both involve people evaluating musical performances. In comparison to other real-life data, the French (presidential) elections seem relatively polarized. Regarding Pabulib (participatory budgeting voting), with few exceptions, the data shows high diversity, with low agreement and low polarization. (This may stem from the fact that, in most cities, people can approve only a very limited number of projects, so the saturation is extremely low). To conclude, it is worth noting that our map of real-life elections resembles the one from the work of Faliszewski et al. [2025b], in which (while focusing on ordinal elections), most real-life instances appeared in a similar region of the map, i.e., low polarization, mid to high diversity, and low to mid agreement.

Finally, we note that MDS embeddings of maps in Figures 2 and 3 achieve mean multiplicative distortions of 1.014 and 1.010, respectively, indicating extremely good representation. In particular, this is far more accurate than the original maps presented in [Szufa et al., 2025, Section 4.3.2], where overall distortion is between 1.21 and 1.26.

²We chose these indices as their sum is the most consistently close to 1 among those studied (see Section C.1 for details), which shows that they are complementary and, hence, together provide a good high-level summary of the nature of a given election. For elections with more than 200 candidates or more than 1000 voters, we computed the indices using sampled subsets of the elections.

8 Future Work

We foresee three directions for future work: First, a deeper study of agreement indices in ordinal elections, going beyond the index we mention in [Section 5.1](#) and Kemeny rule. Second, finding more principled diversity and polarization indices, also for polarization with more than two disagreeing groups. Third, looking for further features of (approval) election and designing indices for them (such as, e.g., measuring if voters approve different candidates independently).

Acknowledgments

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 101002854), from the European Union under the project Robotics and advanced industrial production (reg. no. CZ.02.01.01/00/22_008/0004590), and from the Foundation for the University of Geneva. T. Wąs was supported by the UK Engineering and Physical Sciences Research Council (EPSRC) under grant EP/X038548/1. The research presented in this paper has been partially supported by the funds of Polish Ministry of Science and Higher Education assigned to AGH University of Science and Technology. Further, it was inspired by research proposed for NCN project UMO-2025/58/A/ST6/00371.



References

- Adamic, L. A. and Glance, N. (2005). The political blogosphere and the 2004 US election: Divided they blog. In *Proceedings of LinkKDD-2005*, pages 36–43.
- Alcalde-Unzu, J. and Vorsatz, M. (2013). Measuring the cohesiveness of preferences: An axiomatic analysis. *Social Choice and Welfare*, 41(4):965–988.
- Ammann, M. and Puppe, C. (2025). Preference diversity. *Review of Economic Design*. Online First.
- Baldi, P., Brunak, S., Chauvin, Y., Andersen, C. A. F., and Nielsen, H. (2000). Assessing the accuracy of prediction algorithms for classification: An overview. *Bioinformatics*, 16(5):412–424.
- Bartholdi III, J., Tovey, C., and Trick, M. (1989). Voting schemes for which it can be difficult to tell who won the election. *Social Choice and Welfare*, 6(2):157–165.
- Baujard, A., Igersheim, H., and Delemazure, T. (2025a). Voter autrement 2007 - dataset of the in situ experiments.
- Baujard, A., Igersheim, H., Lebon, I., Gavrel, F., and Delemazure, T. (2025b). Voter autrement 2012 - dataset of the in situ experiments.
- Boehmer, N. (2023). *Application-oriented collective decision making: Experimental toolbox and dynamic environments*. PhD thesis, Technische Universität Berlin.
- Boehmer, N., Brill, M., Cevallos, A., Gehrlein, J., Sánchez-Fernández, L., and Schmidt-Kraepelin, U. (2024). Approval-based committee voting in practice: A case study of (over-)representation in the Polkadot blockchain. In *Proceedings of AAAI-2024*, pages 9519–9527.
- Bouveret, S., Blanch, R., Baujard, A., Durand, F., Igersheim, H., Lang, J., Laruelle, A., Laslier, J.-F., Lebon, I., and Merlin, V. (2018). Voter autrement 2017 - online experiment.
- Can, B., Ozkes, A. I., and Storcken, T. (2015). Measuring polarization in preferences. *Mathematical Social Sciences*, 78:76–79.

- Chierichetti, F., Kumar, R., Pandey, S., and Vassilvitskii, S. (2010). Finding the Jaccard median. In *Proceedings of SODA-2010*, pages 293–311.
- Choueifaty, Y. and Coignard, Y. (2008). Toward maximum diversification. *The Journal of Portfolio Management*, 35(1):40–51.
- Delemazure, T. and Bouveret, S. (2024). Voter autrement 2022 - the online experiment (“un autre vote”).
- Duan, L., Street, W., Liu, Y., Xu, S., and Wu, B. (2014). Selecting the right correlation measure for binary data. *ACM Transactions on Knowledge Discovery from Data*, 9(2):13:1–13:28.
- Eisen, M. B., Spellman, P. T., Brown, P. O., and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. *Proceedings of the National Academy of Sciences*, 95(25):14863–14868.
- Faliszewski, P., Flis, J., Peters, D., Pierczyński, G., Skowron, P., Stolicki, D., Szufa, S., and Talmon, N. (2023a). Participatory budgeting: Data, tools and analysis. In *Proceedings of IJCAI-2023*, pages 2667–2674.
- Faliszewski, P., Janeczko, Ł., Kaczmarczyk, A., Kurdziel, M., Pierczyński, G., and Szufa, S. (2025a). Learning real-life approval elections. In *Proceedings of AAMAS-2025*, pages 704–712.
- Faliszewski, P., Kaczmarczyk, A., Sornat, K., Szufa, S., and Wąs, T. (2023b). Diversity, agreement, and polarization in elections. In *Proceedings of IJCAI-2023*, pages 2684–2692.
- Faliszewski, P., Mertlová, J., Nunn, P., Szufa, S., and Wąs, T. (2025b). Distances between top-truncated elections of different sizes. In *Proceedings of AAAI-2025*, pages 13823–13830.
- Faliszewski, P., Sornat, K., Szufa, S., and Wąs, T. (2026a). Diversity of structured domains via k -Kemeny scores. In *Proceedings of AAAI-2026*, pages 16880–16888.
- Faliszewski, P., Sornat, K., Szufa, S., and Wąs, T. (2026b). Outer diversity of structured domains. In *Proceedings of AAMAS-2026*. To appear. See authors’ website.
- Girvan, M. and Newman, M. E. J. (2002). Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, 99(12):7821–7826.
- Godziszewski, M., Batko, P., Skowron, P., and Faliszewski, P. (2021). An analysis of approval-based committee rules for 2D-Euclidean elections. In *Proceedings of AAAI-2021*, pages 5448–5455.
- Goldberg, K., Roeder, T., Gupta, D., and Perkins, C. (2001). Eigentaste: A constant time collaborative filtering algorithm. *Information Retrieval*, 4(2):133–151.
- Han, X., Lin, L., and Wang, R. (2025). Diversification quotients: Quantifying diversification via risk measures. *Management Science*.
- Harper, F. M. and Konstan, J. A. (2015). The movielens datasets: History and context. *ACM Transactions on Interactive Intelligent Systems*, 5(4):1–19.
- Hashemi, V. and Endriss, U. (2014). Measuring diversity of preferences in a group. In *Proceedings of ECAI-2014*, pages 423–428.
- Hemaspaandra, E., Spakowski, H., and Vogel, J. (2005). The complexity of Kemeny elections. *Theoretical Computer Science*, 349(3):382–391.
- Kamishima, T. (2003). Nantonac collaborative filtering: Recommendation based on order responses. In *Proceedings of KDD-2003*, pages 583–588.
- Karpov, A., Markström, K., Riis, S., and Zhou, B. (2024). Local diversity of Condorcet domains. Technical Report arXiv:2401.11912 [econ.TH], arXiv.org.
- Kemeny, J. (1959). Mathematics without numbers. *Daedalus*, 88:577–591.
- Knuth, D. E. (1993). *The Stanford GraphBase: A platform for combinatorial computing*. ACM Press.

- Krebs, V. E. (2002). Mapping networks of terrorist cells. *Connections*, 24(3):43–52.
- Kruskal, J. (1964). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 29(1):1–27.
- Lackner, M. and Maly, J. (2025). Approval-based shortlisting. *Social Choice and Welfare*, 64(1):97–142.
- Levin, S. A., Milner, H. V., and Perrings, C. (2021). The dynamics of political polarization. *Proceedings of the National Academy of Sciences*, 118(50):e2116950118.
- Lovász, L. and Plummer, M. D. (2009). *Matching Theory*. American Mathematical Society.
- Lusseau, D., Schneider, K., Boisseau, O. J., Haase, P., Slooten, E., and Dawson, S. M. (2003). The bottlenose dolphin community of doubtful sound features a large proportion of long-lasting associations: Can geographic isolation explain this unique trait? *Behavioral ecology and sociobiology*, 54(4):396–405.
- Mattei, N. and Walsh, T. (2013). Preflib: A library for preferences. In *Proceedings of ADT-2013*, pages 259–270.
- Michael, J. H. (1997). Labor dispute reconciliation in a forest products manufacturing facility. *Forest Products Journal*, 47(11/12):41–45.
- Nehring, K. and Puppe, C. (2002). A theory of diversity. *Econometrica*, 70(3):1155–1198.
- Ng, A., Jordan, M., and Weiss, Y. (2001). On spectral clustering: Analysis and an algorithm. In *Proceedings of NeurIPS-2001*, pages 849–856.
- Sokal, R. R. and Michener, C. D. (1958). A statistical method for evaluating systematic relationships. *University of Kansas Scientific Bulletin*, 38(22):1409–1438.
- Szufa, S., Boehmer, N., Brederbeck, R., Faliszewski, P., Niedermeier, R., Skowron, P., Slinko, A., and Talmon, N. (2025). Drawing a map of elections. *Artificial Intelligence*, 343:104332.
- Szufa, S., Faliszewski, P., Janeczko, Ł., Lackner, M., Slinko, A., Sornat, K., and Talmon, N. (2022). How to sample approval elections? In *Proceedings of IJCAI-2022*, pages 496–502.
- Warrens, M. J. (2008). On association coefficients for 2×2 tables and properties that do not depend on the marginal distributions. *Psychometrika*, 73:777–789.
- Xia, L. (2025). A linear theory of multi-winner voting. *arXiv preprint arXiv:2503.03082*.
- Yule, G. U. (1912). On the methods of measuring the association between two attributes. *Journal of the Royal Statistical Society*, 75:579–652.
- Zachary, W. W. (1977). An information flow model for conflict and fission in small groups. *Journal of Anthropological Research*, 33(4):452–473.

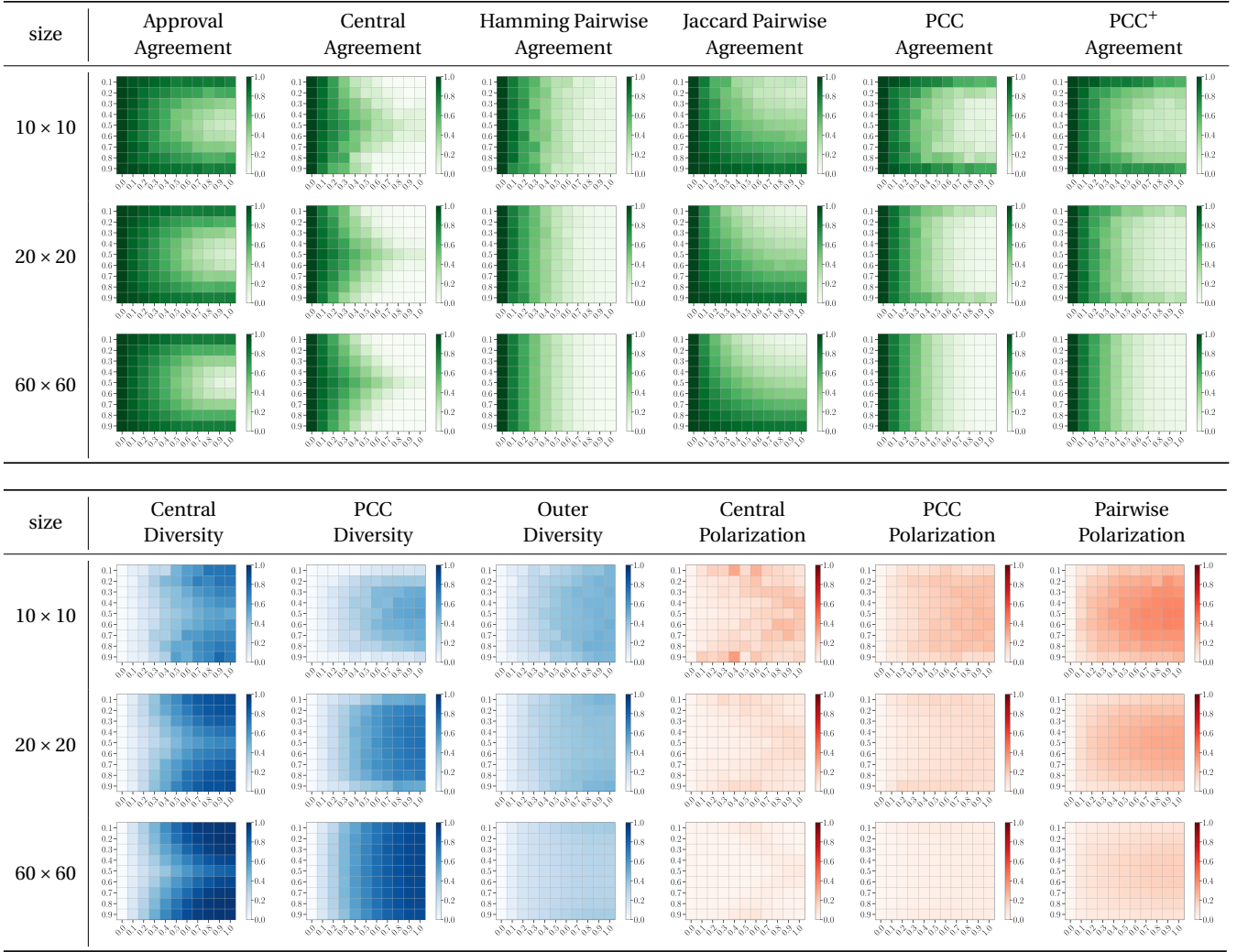


Table 2: Results of the resampling experiment for elections of different sizes.

A Additional Tables

Table 2 shows the results of the resampling experiment for agreement indices and elections of different sizes.

B Missing Proofs

Given two votes, u and v , let n_{11} , n_{10} , n_{01} , and n_{00} be the number of candidates approved in both u and v , only u , only v , and neither of them. Then we have:

$$\text{pcc}(u, v) = \frac{n_{00} \cdot n_{11} - n_{01} \cdot n_{10}}{\sqrt{(n_{10} + n_{11}) \cdot (n_{00} + n_{01}) \cdot (n_{01} + n_{11}) \cdot (n_{00} + n_{10})}}. \quad (2)$$

B.1 Proof of Theorem 5.1

Let $E = (C, V)$ be an election. By definition of the Approval Agreement, we have that:

$$\text{av-agr}(E) = \frac{1}{|C|} \sum_{c \in C} \left| 1 - 2 \frac{|A(c)|}{|V|} \right|.$$

Hence, the value of $\text{av-agr}(E)$ is 1 if and only if for each candidate $c \in C$ it holds that $\left| 1 - 2 \frac{|A(c)|}{|V|} \right| = 1$, which happens exactly if $|A(c)| = 0$ or $|A(c)| = |V|$. This means that $\text{av-agr}(E) = 1$ if and only if E is an identity election. On the other hand, $\text{av-agr}(E) = 0$ if and only if for each candidate $c \in C$ we have $\left| 1 - 2 \frac{|A(c)|}{|V|} \right| = 0$, which happens exactly when $2 \frac{|A(c)|}{|V|} = 1$, i.e., $|A(c)| = |V|/2$. $\square$

B.2 Proof of Theorem 5.2

Let $E = (C, V)$ be an election. By definition of Central Agreement, we have that:

$$\text{cntr-agr}(E) = 1 - \frac{\sum_{v_i \in V} \text{ham}(v_i, u)}{|V| \cdot \min(\text{avl}(E), \text{rev-avl}(E))}.$$

Hence, $\text{cntr-agr}(E)$ is equal to 1 exactly if every summand is equal to 0, which happens when all the votes are identical. Hence, $\text{cntr-agr}(E) = 1$ if and only if E is an identity election.

To characterize when $\text{cntr-agr}(E) = 0$, let us first introduce an auxiliary function. By central Hamming distance (CHD) of election $E = (C, V)$ we mean:

$$\text{chd}(E) = \sum_{v_i \in V} \text{ham}(v_i, u).$$

Naturally, we have:

$$\text{cntr-agr}(E) = 1 - \frac{\text{chd}(E)}{|V| \cdot \min(\text{avl}(E), \text{rev-avl}(E))}.$$

Let us rephrase $\text{chd}(E)$ as follows: Given an election $E = (C, V)$ and one of its central votes $u \in \text{cen}(E)$, for each candidate c_j , we define $\text{chd}(c_j) = \sum_{v_i \in V} |v_i[j] - u[j]|$. Then, by changing the order of summation, we can have:

$$\begin{aligned} \text{chd}(E) &= \sum_{v_i \in V} \text{ham}(v_i, u) \\ &= \sum_{v_i \in V} \sum_{c_j \in C} |v_i[j] - u[j]| \\ &= \sum_{c_j \in C} \sum_{v_i \in V} |v_i[j] - u[j]| = \sum_{c_j \in C} \text{chd}(c_j). \end{aligned} \tag{3}$$

Using this formulation, we characterize the elections with a given size and saturation, for which $\text{chd}(E)$ is largest (specifically, it is easier to verify the claims in the proof of Lemma B.1 by using Equation (3)).

Lemma B.1. *The value $\text{chd}(E)$ is largest among elections with m candidates, n voters, and saturation α if and only if (a) all candidates are approved in some vote in $\text{cen}(E)$, or (b) all candidates are disapproved in some vote in $\text{cen}(E)$, or (c) for each candidate c_j we have that $|A(c_j)| \in \{\lfloor n/2 \rfloor, \lceil n/2 \rceil\}$.*

Proof. Let us first consider case (b). We know that there is an election E with m candidates, n voters, and saturation $\text{satr}(E) = \alpha$, such that all candidates are disapproved in some central vote in $\text{cen}(E)$. We arbitrarily partition the candidates into those that can receive approvals and those that can provide them. Then we perform a sequence of operations where, in each of them, we move a single approval from a providing candidate to a receiving one. We observe that the CHD values remains constant, until the first operation where for the receiving candidate c_k all the central votes for the current election approve it. When this happens, the value of CHD drops by one. From this point on, each operation either maintains the current value of CHD or decreases it further. Since, with an appropriate sequence of operations and an appropriate choice of providing and receiving candidates, we can reach an arbitrary elections with the given size and saturation, it must be the case that the starting election had the maximum value of CHD. The same reasoning applies to case (a), but for moving disapprovals rather than approvals.

Let us consider case (c), assuming that n is odd (if it were even, then cases (a) and (b) would have held as well and we have handled them already). Consider an analogous sequence of operations as in the preceding paragraph, except that the candidates who are originally approved by $\lceil n/2 \rceil$ voters (call them *big* candidates) cannot give approvals to those, who were originally approved by $\lfloor n/2 \rfloor$ voters (call them *small* candidates; this is w.l.o.g., as if we wanted to make such swaps, then we could have simply permuted the candidates). Now we observe that whenever an approval is moved from a small candidate or to a big candidate, then the CHD value of the election decreases. Yet, all our moves are of this form, so the original CHD value must have been maximal.

Note that every election can be reached by moving approvals from one in cases (a), (b), or (c). This completes the proof. $\square$

Next, let us consider some election $E = (C, V)$, such that there is a vote $u \in \text{cen}(E)$ where all candidates are disapproved. By the above lemma, we know that its $\text{chd}(E)$ value is largest for its size and saturation. Further,

we note that:

$$\begin{aligned}\text{chd}(E) &= \sum_{v_i \in V} \text{ham}(v_i, u) \\ &= \sum_{v_i \in V} |A(v_i)| = |V| \cdot \text{avl}(E).\end{aligned}$$

It is also immediate that $|V| \cdot \text{avl}(E) \leq |V| \cdot \text{rev-avl}(E)$ in this case. Similarly, if there is $u \in \text{cen}(E)$ that approves all the candidates, then we have $\text{chd}(E) = |V| \cdot \text{rev-avl}(E)$ and $\text{rev-avl}(E) \leq \text{avl}(E)$. This means that if either all candidates in E are approved by at least half of the voters or all candidates in E are disapproved by at least half of the voters, then $\text{cntr-agr}(E) = 0$.

To finish the proof, it remains to consider an election $E = (C, V)$ where some candidates are approved by fewer than half of the voters (denote the set of these candidates as $C_{<}$), some candidates are approved by more than half of the voters (denote the set of these candidates as $C_{>}$), and some (possibly zero) are approved by exactly half of the voters (denote the set of these candidates as $C_{=}$). Let us form $E_{<} = (C_{<}, V)$, $E_{>} = (C_{>}, V)$ and $E_{=} = (C_{=}, V)$. The following holds (note that $\text{chd}(E_{=}) = |V| \cdot \text{avl}(E_{=}) = |V| \cdot \text{rev-avl}(E_{=})$):

$$\begin{aligned}\text{chd}(E) &= \text{chd}(E_{<}) + \text{chd}(E_{>}) + \text{chd}(E_{=}) \\ &= |V|(\text{avl}(E_{<}) + \text{rev-avl}(E_{>}) + \text{avl}(E_{=}))\end{aligned}$$

One can easily verify that this value is strictly smaller than both $|V| \cdot \text{avl}(E)$ and $|V| \cdot \text{rev-avl}(E)$ (the former inequality holds because $\text{avl}(E) = \text{avl}(E_{<}) + \text{avl}(E_{>}) + \text{avl}(E_{=})$ while $\text{rev-avl}(E_{>}) < \text{avl}(E_{<})$; the latter inequality holds for analogous reasons). Consequently, $\text{cntr-agr}(E) > 0$. $\square$

B.3 Proof of Theorem 5.3

We use the notation as in the proof of Theorem 5.2, in Equation (3). Consequently, we have election $E = (C, V)$ and $\text{chd}(E) = \sum_{c_j \in C} \text{chd}(c_j)$, where:

$$\text{chd}(c_j) = \sum_{v_i \in V} |v_i[j] - u[j]|$$

for some $u \in \text{cen}(E)$. We note that:

$$\text{chd}(c_j) = \begin{cases} |A(c_j)| & \text{if } u[j] = 0, \\ |V| - |A(c_j)| & \text{if } u[j] = 1. \end{cases}$$

Further, we have that $u[j] = 0$ whenever $|A(c_j)| \leq |V|/2$ and $u[j] = 1$ otherwise (strictly speaking, if $|A(c_j)|$ were equal to $|V|/2$ then we could also take $u[j] = 1$ in this case, but this would not change the logic of the argument). Consequently, we have that:

$$\text{chd}(c_j) = \frac{|V|}{2} - \left| \frac{|V|}{2} - |A(c_j)| \right|.$$

After substituting this to the formula for $\text{cntr-agr}(E)$, we get:

$$\begin{aligned}\text{cntr-agr}(E) &= 1 - \frac{\text{chd}(E)}{|V| \cdot \min(\text{avl}(E), \text{rev-avl}(E))} \\ &= 1 - \frac{\sum_{c_j \in C} \left(\frac{|V|}{2} - \left| \frac{|V|}{2} - |A(c_j)| \right| \right)}{|V| \cdot \min(\text{avl}(E), \text{rev-avl}(E))} \\ &= 1 - \frac{\frac{|V|}{2} \sum_{c_j \in C} \left(1 - \left| 1 - \frac{2|A(c_j)|}{|V|} \right| \right)}{|V| \cdot \min(\text{avl}(E), \text{rev-avl}(E))} \\ &= 1 - \frac{\sum_{c_j \in C} \left(1 - \left| 1 - \frac{2|A(c_j)|}{|V|} \right| \right)}{2 \cdot \min(\text{avl}(E), \text{rev-avl}(E))}\end{aligned}$$

This completes the proof. $\square$

B.4 Proof of Theorem 5.4

This result is, in essence, folk knowledge. We provide an argument for the sake of completeness. Let x and y be two vectors of dimension m . We form two new vectors, x' and y' , such that for each $i \in [m]$ we have:

$$x'[i] = \frac{x[i] - \bar{x}[i]}{\sqrt{\sum_{j=1}^m (x[j] - \bar{x}[j])^2}} \text{ and } y'[i] = \frac{y[i] - \bar{y}[i]}{\sqrt{\sum_{j=1}^m (y[j] - \bar{y}[j])^2}}.$$

We observe that $\text{pcc}(x, y)$ is equal to the dot product of x' and y' .

Now let us consider some election $E = (C, V)$, where $V = (v_1, \dots, v_n)$. We form an $n \times n$ matrix M , where entry in the i -th row and j -th column is $\text{pcc}(v_i, v_j)$. By the argument from the previous paragraph, we know that M is a Gram matrix (i.e., a matrix of dot products of a sequence of vectors). Since all Gram matrices are semi-definite positive, and sums of entries of all semi-definite positive matrices are nonnegative, we have that:

$$\sum_{v_i \in V} \sum_{v_j \in V} \text{pcc}(v_i, v_j) \geq 0.$$

On the other hand, since for each two votes v_i and v_j we have that $\text{pcc}(v_i, v_j) \leq 1$, it is immediate that:

$$\sum_{v_i \in V} \sum_{v_j \in V} \text{pcc}(v_i, v_j) \leq n^2.$$

Together, this means that $\text{pcc-agr}(E) \in [0, 1]$. □

B.5 Proof of Theorem 5.5

Let $E = (C, V)$ be an election with candidate set $C = \{c_1, \dots, c_m\}$. We define its expected Hamming distance (EHD) as follows:

$$\text{ehd}(E) = \sum_{u \in V} \sum_{v \in V} \text{ham}(u, v).$$

Naturally, we have:

$$\text{pair-agr}(E) = \text{ehd}(E) / (2|V|^2|C| \cdot \text{satr}(E)(1 - \text{satr}(E))).$$

By definition of $\text{ehd}(E)$ and changing the order of summation, we have:

$$\begin{aligned} \text{ehd}(E) &= \sum_{u \in V} \sum_{v \in V} \text{ham}(u, v) \\ &= \sum_{u \in V} \sum_{v \in V} \sum_{c_j \in C} |u[j] - v[j]| \\ &= \sum_{c_j \in C} \sum_{u \in V} \sum_{v \in V} |u[j] - v[j]|. \end{aligned}$$

Now we note that for each candidate c_j , there are $|A(c_j)|$ voters who approve it and $n - |A(c_j)|$ who do not. Hence, for a given candidate $c_j \in C$, we have that:

$$\begin{aligned} \sum_{u \in V} \sum_{v \in V} |u[j] - v[j]| &= |A(c_j)| \cdot \sum_{v \in V} |1 - v[j]| \\ &\quad + (n - |A(c_j)|) \cdot \sum_{v \in V} |0 - v[j]| \\ &= |A(c_j)|(n - |A(c_j)|) + (n - |A(c_j)|)|A(c_j)| \\ &= 2|A(c_j)|(n - |A(c_j)|). \end{aligned}$$

By substituting this equality into the definition of $\text{ehd}(E)$ we obtain:

$$\text{ehd}(E) = \sum_{c_j \in C} 2|A(c_j)|(n - |A(c_j)|),$$

which substituted to the formula for $\text{pair-agr}(E)$ gives:

$$\text{pair-agr}(E) = \frac{\sum_{c_j \in C} |A(c_j)|(n - |A(c_j)|)}{|V|^2|C| \cdot \text{satr}(E)(1 - \text{satr}(E))}.$$

This completes the proof. □

B.6 Proof of Theorem 5.6

It is easy to verify that, indeed, each of the four indices takes value 1 if and only if the input election is an identity one. Further, closer inspection of the definition of Jaccard Pairwise Agreement indicates that it assumes value 0 if and only if all voters approve disjoint sets of candidates (otherwise one of the summands in its definition would be greater than 0). It remains to consider when Hamming Pairwise Agreement assumes value 0.

We use the same notation as in the proof of Theorem 5.5 and we first prove the next lemma, which characterizes for which elections E of a given size and saturation value $\text{ehd}(E)$ is largest.

Lemma B.2. Let m and n be two positive integers, and let $\alpha \in [0, 1]$ be a rational number such that $\alpha \cdot nm$ is an integer. Let $E = (C, V)$ be an election with m candidates, n voters, and $\text{satr}(E) = \alpha$. Value $\text{ehd}(E)$ is largest among elections of its size and saturation if for each two candidates $c_j, c_k \in C$, $||A(c_j)| - |A(c_k)|| \leq 1$.

Proof. Let $E = (C, V)$ be an arbitrary election such that $C = \{c_1, \dots, c_m\}$, $V = (v_1, \dots, v_n)$, and $\text{satr}(E) = \alpha$. We claim that if there are two candidates $c_j, c_k \in C$ such that c_k has at least two more approvals than c_j then by swapping two approvals within E we can obtain a new election E' such that $\text{ehd}(E') > \text{ehd}(E)$.

Let us first express $\text{ehd}(E)$ in a convenient way. Let us define function $f(x) = x(1-x)$ and for each candidate $c_t \in C$, let p_t be its approval probability, i.e., $p_t = |A(c_t)|/n$. Now, using the formula for $\text{ehd}(E)$ given in the proof of [Theorem 5.5](#), we have:

$$\begin{aligned} \text{ehd}(E) &= \frac{2}{n^2} \sum_{c_t \in C} n \cdot f(p_t) \\ &= \frac{2}{n} \sum_{c_t \in C} f(p_t) \\ &= \frac{2}{n} (f(p_j) + f(p_k)) + X(E), \end{aligned}$$

where $X(E) = \frac{2}{n} \sum_{c_t \in C \setminus \{c_j, c_k\}} f(p_t)$. Next, we form election E' that is identical to E except that we remove a single approval from c_k and add a single approval to c_j . We have that:

$$\text{ehd}(E') = \frac{2}{n} (f(p_j + 1/n) + f(p_k - 1/n)) + X(E),$$

To establish that $\text{ehd}(E') > \text{ehd}(E)$, we need to show that

$$f(p_j + 1/n) - f(p_j) > f(p_k) - f(p_k - 1/n). \quad (4)$$

If we have that both $p_j, p_k \leq 0.5$ or both $p_j, p_k \geq 0.5$ then it is clear that this holds (as the derivative of $f(x)$ is strictly decreasing and assumes value 0 for $x = 0.5$). On the other hand, if $p_j < 0.5$ and $p_k > 0.5$ then either the left-hand side of (4) is positive and the right-hand side is at most zero, or the left-hand side is zero and the right-hand side is negative (this follows from the shape of function $f(x)$, which grows for $x \in [0, 0.5]$ and then symmetrically decreases for $x \in [0.5, 1]$, and the fact that c_k has at least two more approvals than c_j in E).

By repeating the above reasoning, we find that elections where the numbers of approvals received by some two candidates differ by more than one cannot have the largest EHD value. By [Theorem 5.5](#) we know that an election's EHD value—for a given number of voters—depends only on the candidates' approval scores. Consequently, an election has the highest EHD value among those with a given numbers of candidates and voters, and a given saturation, if and only if approval scores of its candidates differ at most by one. $\square$

Now let us consider an election $E = (C, V)$, where each candidate has the same approval score equal to T . In such an election, we have $\text{satr}(E) = T/|V|$. By [Theorem 5.5](#) and substituting candidates' approval scores into the formula, we have that:

$$\begin{aligned} \text{pair-agr}(E) &= 1 - \frac{|C|T(|V|-T)}{|V|^2|C|\text{satr}(E)(1-\text{satr}(E))} \\ &= 1 - \frac{|C|T(|V|-T)}{|V|^2|C| \cdot T/|V|(1-T/|V|)} \\ &= 1 - \frac{|C|T(|V|-T)}{T(|V|-T)} = 0. \end{aligned}$$

Next, we observe that for an election where candidates' approval scores differ by at most one (and, indeed, there are two candidates with different scores) it holds that $\text{pair-agr}(E) > 0$. To this end, let us consider election $E = (C, V)$ with m candidates and n voters, where T voters approve all the candidates, 1 voter approves $X < m$ candidates, and remaining $n - T - 1$ voters do not approve any candidates. We have:

$$\begin{aligned} \text{ehd}(E) &= T(0 + (m - X) + m(n - T - 1)) \\ &\quad + T(m - X) + 0 + X(n - T - 1) \\ &\quad + (n - T - 1)(mT + X + 0) \\ &= 2T(m - X) + 2mT(n - T - 1) + 2X(n - T - 1) \\ &= 2mT - 2TX + 2nmT - 2mT^2 \\ &\quad - 2mT + 2Xn - 2TX - 2X \\ &= 2nmT - 4TX - 2mT^2 + 2Xn - 2X \\ &= n^2m \left(\frac{2T}{n} - \frac{4TX}{n^2m} - \frac{2T^2}{n^2} + \frac{2X}{nm} - \frac{2X}{n^2m} \right). \end{aligned}$$

For election E we have that $\text{satr}(E) = \frac{mT+X}{nm}$, so:

$$\begin{aligned} \text{pair-agr}(E) &= 1 - \frac{n^2 m (\frac{2T}{n} - \frac{4TX}{n^2 m} - \frac{2T^2}{n^2} + \frac{2X}{nm} - \frac{2X}{n^2 m})}{2n^2 m (\frac{mT+X}{nm})(1 - \frac{mT+X}{nm})} \\ &= 1 - \frac{\frac{2T}{n} - \frac{4TX}{n^2 m} - \frac{2T^2}{n^2} + \frac{2X}{nm} - \frac{2X}{n^2 m}}{2(\frac{mT+X}{nm})(1 - \frac{mT+X}{nm})} \\ &= 1 - \frac{\frac{2T}{n} - \frac{4TX}{n^2 m} - \frac{2T^2}{n^2} + \frac{2X}{nm} - \frac{2X}{n^2 m}}{\frac{2T}{n} - \frac{4TX}{n^2 m} - \frac{2T^2}{n^2} + \frac{2X}{nm} - \frac{2X^2}{n^2 m^2}} > 0. \end{aligned}$$

The final inequality follows by computing the numerator and denominator. This concludes the proof as, by [Theorem B.2](#), for every other election of this size and saturation, $\text{pair-agr}(E)$ is even larger.

B.7 Proof of Theorem 5.7

Let $E = (C, V)$ be a k -Party election. We write m and n to denote $|C|$ and $|V|$, and we assume that both m and n are divisible by k . Given two votes, u and v in V , that approve the same groups of m/k candidates, we have that:

$$\text{pcc}(u, v) = 1.$$

On the other hand, if u and v approve disjoint sets of m/k candidates each, then, by Equation (2) we have:

$$\text{pcc}(u, v) = \frac{-\frac{m}{k} \cdot \frac{m}{k}}{\sqrt{\frac{m}{k} \cdot (m - \frac{m}{k}) \cdot \frac{m}{k} \cdot (m - \frac{m}{k})}} = -\frac{\frac{m^2}{k^2}}{\frac{m^2}{k} (1 - \frac{1}{k})} = -\frac{1}{k} \frac{1}{1 - \frac{1}{k}}.$$

Then, we have that (the second equality follows by substituting the above values for PCCs):

$$\begin{aligned} \text{pcc-agr}(E) &= \frac{1}{n^2} \sum_{v_i \in V} \sum_{v_j \in V} \text{pcc}(v_i, v_j) \\ &= \frac{1}{n^2} \cdot n \cdot \left(\frac{n}{k} - (n - \frac{n}{k}) \frac{1}{k} \frac{1}{1 - \frac{1}{k}} \right) \\ &= \left(\frac{1}{k} - \frac{1}{k} \right) = 0. \end{aligned}$$

Similarly, we have:

$$\begin{aligned} \text{pcc}^+ \text{-agr}(E) &= \frac{1}{n^2} \sum_{v_i \in V} \sum_{v_j \in V} \max(0, \text{pcc}(v_i, v_j)) \\ &= \frac{1}{n^2} \cdot n \cdot \frac{n}{k} = \frac{1}{k}. \end{aligned}$$

This completes the proof. □

B.8 Proof of Theorem 5.8

Let $E = (C, V)$ be an election such that $0 < |A(u)| = |A(v)| < 1$ for all $u, v \in V$.

First, we derive an explicit dependence between the $\text{pcc}(u, v)$ and $\text{ham}(u, v)$ for any pair of $u, v \in V$. We note that an implicit dependence has been known in the literature on association coefficients (see, e.g., [Warrens \[2008\]](#)); for completeness we provide the exact formula in our notation.

Specifically, [Warrens \[2008\]](#) considers Yule's ϕ coefficient (denoted there by S_{Yule1}) [Yule \[1912\]](#), which coincides with PCC for binary vectors, and also the simple matching coefficient S_{SM} [Sokal and Michener \[1958\]](#), which in our notation equals to the number of matches, i.e., $S_{\text{SM}}(u, v) = m - \text{ham}(u, v)$. [Warrens \[2008\]](#) shows that, under fixed marginals, S_{Yule1} can be expressed as an affine transformation of S_{SM} , but do not provide an explicit formula, which is as follows.

Proposition B.3. *If $0 < |A(u)| = |A(v)| = pm < m$ holds then we have $\text{pcc}(u, v) = 1 - \frac{\text{ham}(u, v)}{2m(1-p)p}$.*

Proof. Let $n_{11}, n_{10}, n_{01}, n_{00}$ denote the numbers of candidates approved in both u and v , only in u , only in v , and in neither of them, respectively. By these definitions we have

$$n_{11} + n_{10} = n_{11} + n_{01} = pm. \tag{5}$$

This implies:

$$\begin{aligned}
n_{00} &= m - n_{11} - n_{10} - n_{01} \\
&\stackrel{(5)}{=} m - n_{11} - (pm - n_{11}) - (pm - n_{11}) \\
&= m - 2pm + n_{11}.
\end{aligned}$$

Now consider the numerator of $\text{pcc}(u, v)$ (defined in Equation (2)).

$$\begin{aligned}
&n_{11} \cdot n_{00} - n_{01} \cdot n_{10} \\
&\stackrel{(5)}{=} n_{11} \cdot (m - 2pm + n_{11}) - (pm - n_{11})^2 \\
&= n_{11} \cdot m - (pm)^2.
\end{aligned}$$

The denominator of $\text{pcc}(u, v)$ equals to:

$$\begin{aligned}
&\sqrt{(n_{10} + n_{11}) \cdot (n_{00} + n_{01}) \cdot (n_{01} + n_{11}) \cdot (n_{00} + n_{10})} \\
&= \sqrt{pm \cdot (m - pm) \cdot pm \cdot (m - pm)} \\
&= pm \cdot (m - pm) = m^2 \cdot p(1 - p).
\end{aligned}$$

Putting both together we obtain:

$$\begin{aligned}
\text{pcc}(u, v) &= \frac{n_{11} \cdot m - (mp)^2}{m^2 \cdot p(1 - p)} \\
&= \frac{n_{11} - mp^2}{mp(1 - p)}.
\end{aligned} \tag{6}$$

By the definition of the Hamming distance we have $\text{ham}(u, v) = n_{10} + n_{01} = 2(pm - n_{11})$ which is equivalent to

$$n_{11} = pm - \frac{\text{ham}(u, v)}{2}. \tag{7}$$

Putting these together we obtain the final equality:

$$\begin{aligned}
\text{pcc}(u, v) &\stackrel{(6),(7)}{=} \frac{mp - \frac{\text{ham}(u, v)}{2} - mp^2}{mp(1 - p)} \\
&= \frac{mp - mp^2}{mp(1 - p)} - \frac{\frac{\text{ham}(u, v)}{2}}{mp(1 - p)} \\
&= 1 - \frac{\text{ham}(u, v)}{2mp(1 - p)}.
\end{aligned}$$

□

Since $|A(v)| = pm$ for all $v \in V$, we have $\text{avl}(E) = pm$, and thus $\text{satr}(E) = \frac{\text{avl}(E)}{m} = p$ and together with Theorem B.3 we can prove the final equality:

$$\begin{aligned}
\text{pcc-agr}(E) &\stackrel{\text{Def. 5.3}}{=} \frac{1}{n^2} \sum_{u \in V} \sum_{v \in V} \text{pcc}(u, v) \\
&\stackrel{\text{Prop. B.3}}{=} \frac{1}{n^2} \sum_{u \in V} \sum_{v \in V} \left(1 - \frac{\text{ham}(u, v)}{2mp(1 - p)} \right) \\
&= 1 - \sum_{u \in V} \sum_{v \in V} \frac{\text{ham}(u, v)}{2n^2 mp(1 - p)} \\
&\stackrel{\text{Def. 5.3}}{=} \text{pair-agr}(E).
\end{aligned}$$

This completes the proof.

□

B.9 Proof of Theorem 6.1

Proof. We will prove that both sides of the equation can be linked to the ratios of the sizes of two combinatorial sets defined as follows. Note that in $p\text{-}\mathcal{U}_C$ every vote appears $a^r \cdot b^{|C|-r}$ times, where a and b are some integers such that $p = a/(a+b)$ and r is the number of ones in the vote. Let F be a set of functions from the set of m numbers to the set of $a+b$ numbers, $f: [m] \rightarrow [a+b]$. Then, let X be set of pairs $(x, f) \in [m] \times F$, such that $f(x) \in [a]$, if $x \in [(1-q)m]$, or $f(x) \in \{a+1, \dots, a+b\}$, if $x \in \{(1-q)m+1, \dots, m\}$. We claim that (1) $\text{ham}(\{v\}, p\text{-}\mathcal{U}_C) = |X|/(m \cdot |F|)$ and (2) $|X|/(m \cdot |F|) = p(1-q) + q(1-p)$.

(1) First, observe that

$$|F| = (a+b)^m = \sum_{k=0}^m \binom{m}{k} a^k b^{m-k} = |p\text{-}\mathcal{U}_C|. \quad (8)$$

Now, without loss of generality, let us assume that the first $(1-q)m$ positions of v are 0s and the last qm positions are 1s. According to the definition of $p\text{-}\mathcal{U}_C$, each vote $u \in \{0,1\}^m$ appears there $a^\alpha b^\beta$ times, where α is the number of ones in u and β is the number of zeros. This corresponds to the number of functions f that give values in $[a]$ to exactly those arguments in $[m]$ that belong to $A(u)$. Now, for each such function f , we can choose x in the number of different ways that is equal to the number of indices in $A(u)$ that are also in $[(1-q)m]$ plus the number of indices that are not in $A(u)$ but are in $\{(1-q)m+1, \dots, m\}$, i.e., the symmetric difference between $A(u)$ and $\{(1-q)m+1, \dots, m\}$. But this is exactly the Hamming distance between u and v . Hence, $|X|$ is equal to the sum of the (unnormalized) Hamming distances between v and each vote from $p\text{-}\mathcal{U}_C$. Combining this with eq. (8), we get that indeed $\text{ham}(\{v\}, p\text{-}\mathcal{U}_C) = |X|/(m \cdot |F|)$.

(2) To express the size of X in a different way, we first fix $x \in [m]$, and then select matching f . We split X into two disjoint subsets depending on the value of x , in a way that $X_1 = \{(x, f) \in X : x \in [(1-q)m]\}$ and $X_2 = \{(x, f) \in X : x \in \{(1-q)m+1, \dots, m\}\}$.

In X_1 we can select x in $(1-q)m$ different ways, then $f(x)$ can have a different values, and $f(y)$ for each $y \in [m] \setminus \{x\}$ can have $(a+b)$ different values. Thus, $|X_1| = (1-q)m \cdot a \cdot (a+b)^{m-1}$. Analogously, $|X_2| = q \cdot m \cdot b \cdot (a+b)^{m-1}$. Since X_1 and X_2 are disjoint and $X = X_1 \cup X_2$, we get that $|X| = |X_1| + |X_2| = ((1-q)a + qb)m(a+b)^{m-1}$. Thus, by eq. (8)

$$\frac{|X|}{m|F|} = \frac{(1-q)a + qb}{a+b} = (1-q)p + q(1-p).$$

□

B.10 Proof of Theorem 6.2

Proof. Clearly, for every election $E = (C, V)$ and $p = \text{satr}(E)$, it holds that $\text{ham}(V, p\text{-}\mathcal{U}_C) \geq 0$, hence $\text{out-div}_h^{\text{satr}}(E) \leq 1$. For the lower bound, by a series of (in)equalities, we will show that $\text{ham}(V, p\text{-}\mathcal{U}_C) \leq 2(p - p^2)$, which will imply that $\text{out-div}_h^{\text{satr}}(E) \geq 0$.

Denote $V = \{v_1, \dots, v_n\}$. Let us start by considering copying n times each vote in $p\text{-}\mathcal{U}_C$, we denote the resulting collection of votes as $n \cdot p\text{-}\mathcal{U}_C$. Observe that introducing these additional copies does not affect the Hamming distance from V (as otherwise it would mean that there is a fractional matching giving less total weight than an integral matching, which is not possible [Lovász and Plummer \[2009\]](#)). Hence,

$$\text{ham}(V, p\text{-}\mathcal{U}_C) = \text{ham}(V, n \cdot p\text{-}\mathcal{U}_C).$$

To compute $\text{ham}(V, n \cdot p\text{-}\mathcal{U}_C)$ we match the votes from V to those in $n \cdot p\text{-}\mathcal{U}_C$ in an optimal way. Another way of matching the votes would be to match each separate set of copies of votes in $p\text{-}\mathcal{U}_C$ to a single, each time different vote in V . Thus, we get

$$\begin{aligned} \text{ham}(V, p\text{-}\mathcal{U}_C) &= \text{ham}(V, n \cdot p\text{-}\mathcal{U}_C) \\ &\leq \frac{1}{n} \sum_{i=1}^n \text{ham}(\{v_i\}, p\text{-}\mathcal{U}_C). \end{aligned}$$

Now, let us denote a fraction of ones in v_i by $q_i = |A(v_i)|/|C|$. From [Theorem 6.1](#), we get that

$$\begin{aligned} \text{ham}(V, p\text{-}\mathcal{U}_C) &\leq \frac{1}{n} \sum_{i=1}^n \text{ham}(\{v_i\}, p\text{-}\mathcal{U}_C) \\ &= \frac{1}{n} \sum_{i=1}^n (p \cdot (1 - q_i) + (1 - p) \cdot q_i) \\ &= p \left(1 - \frac{1}{n} \sum_{i=1}^n q_i \right) + (1 - p) \cdot \frac{1}{n} \sum_{i=1}^n q_i \\ &= 2p(1 - p), \end{aligned}$$

where the last equality we can obtain from the fact that $\sum_{i=1}^n q_i = \text{satr}(E) = p$. This concludes the proof. $\square$

C Further Data Analysis

In this appendix, we elaborate on the measure of complementarity of computed indices, and provide correlation matrix of their values.

C.1 Complementarity

[Faliszewski et al. \[2023b\]](#) suggest that the elections they consider approximately fit into a simplex in the space of their agreement, diversity, and polarization indices. We would like to measure how much this is the case for all triples of our indices. To this end, we use the following *measure of complementarity* (we note that similar measures have been used to measure a diversification of a portfolio in financial sciences, see e.g., [Choueifaty and Coignard \[2008\]](#); [Han et al. \[2025\]](#)).

Definition C.1. Let $\mathbf{x} = (x_1, \dots, x_n)$, $\mathbf{y} = (y_1, \dots, y_n)$, and $\mathbf{z} = (z_1, \dots, z_n)$, be three vectors of index values. We define their complementarity as one minus the standard deviation of the sum of the indices divided by the sum of the standard deviation of the indices, i.e.,

$$\text{cmpl}(\mathbf{x}, \mathbf{y}, \mathbf{z}) = 1 - \frac{\text{std}(x_1 + y_1 + z_1, \dots, x_n + y_n + z_n)}{\text{std}(\mathbf{x}) + \text{std}(\mathbf{y}) + \text{std}(\mathbf{z})},$$

where

$$\text{std}(a_1, \dots, a_n) = \sqrt{\frac{1}{n} \sum_{i=1}^n (a_i - \frac{1}{n} \sum_{j=1}^n a_j)^2}.$$

Intuitively, complementarity measures how close is the sum of indices to being constant as compared to the variability of the indices alone. If the sum is constant, complementarity would give value equal to 1. On the other hand, for any vector $\mathbf{x} \in \mathbb{R}^n$, we have $\text{cmpl}(\mathbf{x}, \mathbf{x}, \mathbf{x}) = 0$.

We checked the values of the complementarity measure for all triples of our indices. The four that gave the highest values where:

- pcc-agr, pcc-div, pcc-pol: the value of 0.9016,
- cntr-agr, cntr-div, pcc-pol: the value of 0.8793,
- cntr-agr, cntr-div, cntr-pol: the value of 0.8775, and
- pair-agr, cntr-div, pcc-pol: the value of 0.8748.

This motivated us to focus on the pcc-agr, pcc-div, pcc-pol map in [Section 7](#).

For comparison, in the original paper by [Faliszewski et al. \[2023b\]](#) that introduced diversity-agreement-polarization framework, the complementarity of their indices for the ordinal elections in their *extended dataset* is equal to 0.8818 (we have normalized the values of the indices to have the maximum value of 1, as it is not the case there). In the follow-up paper by [Faliszewski et al. \[2025b\]](#), the complementarity of the indices of the election in the main dataset is equal to 0.9086.

	satr	av-agr	cntr-agr	pair-agr	pcc-agr	pcc ⁺ -agr	jacc-agr	cntr-div	pcc-div	out-div	cntr-pol	pcc-pol	pair-pol
satr	–	-0.496	0.425	0.163	0.032	0.115	0.644	-0.345	-0.121	-0.111	0.132	0.149	0.294
app_agr	-0.461	–	-0.143	0.177	0.320	0.100	-0.202	0.211	-0.057	0.049	-0.247	-0.302	-0.490
cntr_agr	0.374	0.095	–	0.554	0.399	0.350	0.580	-0.567	-0.328	-0.251	0.094	0.100	0.196
pair_agr	0.243	0.328	0.909	–	0.709	0.453	0.397	-0.437	-0.413	-0.157	0.080	0.042	0.044
pcc_agr	0.106	0.480	0.690	0.816	–	0.650	0.281	-0.288	-0.531	-0.163	0.009	-0.052	-0.007
pcc ⁺ _agr	0.173	0.242	0.613	0.691	0.855	–	0.422	-0.555	-0.827	-0.421	0.279	0.270	0.302
jacc_agr	0.812	-0.127	0.689	0.656	0.516	0.603	–	-0.592	-0.421	-0.318	0.219	0.195	0.287
cntr_div	-0.392	0.198	-0.668	-0.626	-0.420	-0.699	-0.707	–	0.636	0.469	-0.477	-0.490	-0.464
pcc_div	-0.174	-0.115	-0.495	-0.576	-0.667	-0.913	-0.577	0.832	–	0.498	-0.366	-0.407	-0.363
out_div	-0.134	0.040	-0.374	-0.371	-0.295	-0.503	-0.398	0.642	0.609	–	-0.222	-0.252	-0.280
cntr_pol	0.181	-0.379	-0.059	-0.059	-0.110	0.353	0.284	-0.649	-0.567	-0.461	–	0.613	0.497
pcc_pol	0.180	-0.399	-0.028	-0.081	-0.151	0.354	0.276	-0.669	-0.604	-0.504	0.929	–	0.619
pair_pol	0.257	-0.516	0.030	-0.081	-0.125	0.360	0.295	-0.640	-0.569	-0.444	0.817	0.912	–

Table 3: Correlations between our features based on the values for elections in the main dataset. The values in the top-right triangle are Kendall’s τ correlation coefficients, while those in the bottom-left triangle are Pearson’s ρ correlation coefficients.

C.2 Correlations

In Table 3 we report the Pearson’s ρ and Kendall’s τ correlation coefficients for each pair of our indices.

We observe that agreement indices (except for Approval Agreement), diversity indices, and polarization indices all form relatively well correlated clusters which confirms our intuitions that they are measuring more or less the same features. Across these clusters there is usually either negative correlation or a lack of correlation. The only exception is the fact that Jaccard and PCC⁺ Pairwise Agreement indices are slightly correlated with the polarization indices. This is due to the fact that Jaccard and PCC⁺ Pairwise Agreement indices capture the local agreement, which means that the presence of voter groups that have similar preference inside each group, but possibly very dissimilar across the groups, increases the value of these two agreement indices. However, the same structures also increase the polarization indices.

Finally, we note that the Saturation index is not significantly correlated with any other index, except for Jaccard Pairwise Agreement. This can be explained by noting that these are the only two indices that we consider that do not treat approvals and disapprovals symmetrically.

D Detailed Datasets Description

In this appendix we provide a complete description of the data used to produce Figures 2 and 3. We begin by providing details on exact distributions used to generate synthetic elections as well as sources of real-world data and the methods of its preprocessing. We conclude with Table 4, in which for each election in the dataset we provide the values of all considered features as well as its icon akin to those in Table 1.

D.1 Synthetic Data

Let us start by describing elections generated by us using synthetic distributions.

Compass Compass elections include exactly those that are listed in Table 1, one copy each.

Impartial Culture We generated ten IC elections with 100 candidates and 1000 voters, two for each value of $p = \{0.1, 0.3, 0.5, 0.7, 0.9\}$.

Lin-IC We generated five Lin-IC elections with 100 candidates and 1000 voters each, as described in Section 3.

Resampling We generated 50 resampling elections drawing ϕ and p uniformly at random from $[0, 1]$, each with 100 candidates and 1000 voters.

N(2-Party) We generated 25 noisy 2-party elections, each with 100 voters and 100 candidates. We sampled ϕ uniformly at random from $[0, 1]$.

(x, y) -2-Party We generated 25 (x, y) -2-Party elections, with 100 voters and 100 candidates each. For each we chose x and y uniformly at random and independently from $[0, 1]$.

Party-List We generated 25 uneven party-list elections, each with 100 voters and 100 candidates, in the following way. First we sampled the number of parties k from a Poisson distribution with $\lambda = 1$ and added 3, so that we have at least 3 parties. Next, we uniformly at random chose two partitions of 100 into k numbers. One of the partitions was then used as the number of voters in each party, while the other was used as the number of candidates.

ID Mixture For $p \in [0, 1]$ and integer k , in k, p -ID Mixture model, we first split the voters randomly into k groups. Then, each group has identical approvals drawn from p -IC. We generated 20 ID Mixture elections, each with 100 candidates and 1000 voters, one for each combination of $k \in \{2, 3, 4, 5\}$ and $p \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$.

IAM Mixture For a given k , we generate an IAM Mixture election by first splitting the voters randomly into k groups, then for each group and candidate c drawing p_c uniformly at random from $[0, 1]$, and finally sampling a sub-election for each group using the Independent Approval Model described in Section 3. We generated 20 IAM Mixture elections, each with 100 candidates and 1000 voters, five for each $k \in \{2, 3, 4, 5\}$.

2D-Euclidean We sampled 50 election from the 2D Euclidean model, replicating first five distributions from Godziszewski et al. [2021] with 10 elections in each (each with 100 candidates and 1000 voters). Specifically, we assumed that both voters and candidates are sampled uniformly from $[0, 1] \times [0, 1]$ square, and then:

- each voter i approves all candidates in radius $r = 0.117$ from i 's position,
- each voter i approves all candidates in radius $r = 0.167$ from i 's position,
- each voter i approves all candidates in radius r_i from i 's position, where r_i is drawn uniformly from $[0, 0.5]$,
- each voter i approves 10 closest candidates, or
- each voter i approves some number of closest candidates, that is drawn uniformly from $\{1, 2, \dots, 100\}$.

D.2 Real-World Elections

Let us now cover elections that are based on real-world data. We include elections from Preflib Mattei and Walsh [2013], Pabulib Faliszewski et al. [2023a], as well as a number of datasets that we have gathered or converted ourselves.

Chopin These elections are based on the XIX International Chopin Piano Competition. Each election is based on one of four stages of the competition. In each, every member of the jury gives a score from 1 to 25 to every participant of a given stage (except for a relatively rare case in which the participant is a student of the member of the jury). The jury is fixed and consists of 17 members, which are voters in our election, and the number of participants, i.e., candidates, decreases with each stage (starting from 84 in the first stage and ending at 11 in the final stage) as only the best candidates pass to the next stage. For each stage, we converted a score above 20 to an approval.

French Elections Each of these datasets consists of exit-poll surveys conducted during the French presidential elections of 2007 Baujard et al. [2025a], 2012 Baujard et al. [2025b], 2017 Bouveret et al. [2018], and 2022 Delemazure and Bouveret [2024].

Jester The Jester approval elections were derived from a Kaggle dataset³ in which users assigned numerical scores to jokes [Goldberg et al. \[2001\]](#). In order to convert the data into approval profile, jokes were ranked according to their scores, with unranked jokes placed randomly at the bottom of the ranking. A fixed top percentage of jokes was then labeled as approved candidates, while the remaining jokes were considered unapproved.

MovieLens This dataset is a subset of MovieLens 20M Dataset that can be found on Kaggle [Harper and Konstan \[2015\]](#).⁴ The subset consists of 50 films (candidates) and 423 users (voters) each of which rates each film with a score from $\{1.0, 1.5, 2.0, \dots, 5.0\}$. We have converted this subset into three elections by setting an approval for every score that is equal or above 4.0, 4.5, or 5.0.

Networks We have included seven elections that have been converted from canonical network datasets. In each of them we have treated the set of nodes as both voters and candidates with an approval corresponding to an edge from a voter to a candidate. Additionally, each node approves itself.

Three networks have been taken from the Mark Newman's collection of network datasets.⁵ These include American College Football Network [Girvan and Newman \[2002\]](#), Dolphin Social Network [Lusseau et al. \[2003\]](#), and Political Blogs Network [Adamic and Glance \[2005\]](#). The next four network has been taken from the predefined datasets in the Editor4Centralities online tool.⁶ These include September 11 Attack Terrorist Network [Krebs \[2002\]](#), Wood-Processing Facility Strike Network [Michael \[1997\]](#), Zachary's Karate Club Network [Zachary \[1977\]](#), and Les Miserables Character Network [Knuth \[1993\]](#).

Polkadot We sampled 10 elections from the Polkadot dataset on Preflib. Each election represents votes of the stakeholders in the Polkadot blockchain over the set of validators in the Nominated Proof-of-Stake (NPoS) protocol [Boehmer et al. \[2024\]](#).

Our Surveys We collected these datasets in an online survey filled predominantly by students. Voters were asked to mark their like or dislike for each of the candidates which then translated to approvals and disapprovals. No restrictions regarding the vote length were imposed. The topics of the surveys include colors, fruits, pets, pizza types, and school subjects (each topic corresponds to one election).

Bidding Data We include six elections based on bids of reviewers over the papers which they want to review in computer science conferences. All of the data comes from Preflib. Three elections comes from AAMAS 2015, 2016, and 2021 editions and three come from unspecified conferences. In each case, we have interpreted "yes" bids as an approval, and all other responses (and lack thereof) as a lack of approval.

Campsongs These four elections are surveys in which the participants had to state their preferences over CCM-genre music pieces. The data is available on Preflib.

CTU This is an approval election over the time of tutorial at the Czech Technical University in Prague.

Eurovision We include a sample of 20 election based on the Eurovision Song Contest from the dataset available on Preflib [Boehmer \[2023\]](#). Originally the data consists of rankings over most preferred participants as seen by each member of the jury (the number of ranked participants is fixed each time). We have converted these top-truncated ordinal elections to approval elections, by setting an approval each for each ranked participant.

Posters These are two approval elections from the San Sebastian Poster Competition available at Preflib. Two elections were held over two groups of posters (A group) and (B group).

³<https://www.kaggle.com/datasets/aakaashjois/jester-collaborative-filtering-dataset>

⁴<https://www.kaggle.com/datasets/grouplens/movielens-20m-dataset>

⁵<https://websites.umich.edu/~mejn/netdata/>

⁶<https://centrality.mimuw.edu.pl/editor/>

Sushi These are two elections based on two surveys on the favorite kinds of sushi [Kamishima \[2003\]](#). In one of them the participants had to indicate most preferred 10 types from the total of 100 types and provide their strict ranking. We have converted this data into an approval election by setting an approval for each of the selected 10 types. In the other survey the participants were allowed to provide a weak ordering of the 10 chosen types (with ties). Here, we included only the kinds that were tied at the first place.

Pol.is The data comes from online deliberation platform Pol.is, in which the users can express their support, neutrality, or disapproval over the comments posted by other users. The data is incomplete, as not every user has to react to every comment. We have transformed all datasets of this type available at Preflib into approval elections by recognizing the support as an approval.

Pabulib Finally, we include a sample of participatory budgeting elections from Pabulib [Faliszewski et al. \[2023a\]](#). In each such election, citizens of a given municipality or district might express their preferences over the projects that can be implemented in their neighborhood. Most of the elections that we included are natively approval elections (the only exception is Toulouse 2019 election in which the voters could distribute 7 points between the projects—we treated each positive amount of points as an approval). However, the projects are associated with costs that could impact the voters decisions, and which are ignored for our purposes. We include five elections from each of the following five cities: Amsterdam, Budapest, Łódź, Toulouse, Warsaw. The elections included are the ones that have the highest “quality” score on the Pabulib webpage. Apart from that, we included all elections from the districts of Warsaw in 2026.

D.3 Index Values

The following is the table with values of our indices for all elections in our main dataset. For each election we present an icon representing its approval matrix, akin to those presented in [Table 1](#). If the election had more than 500 candidates (or voters), we have sampled 500 candidates (or voters) for the purpose of drawing the icon. For most of the elections we ordered the rows (and columns) in the icon based on the Jaccard distances between the votes (and candidates). Based on the distances we computed the tree of hierarchical clustering and ordered the rows (and columns) according to the ordering of the leaves in the tree [Eisen et al. \[1998\]](#).

Election	Icon	C	V	satr	av-agr	cntr-agr	pair-agr	pcc-agr	pcc+agr	jacc-agr	cntr-div	pec-div	out-div	cntr-pol	pec-pol	pair-pol
1/3-ID		60	60	0.33	1.00	1.00	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00
2-Party		60	60	0.50	0.00	0.00	0.00	0.00	0.50	0.50	0.20	0.23	0.10	1.00	1.00	1.00
N(2-Party, 0.6)		60	60	0.49	0.08	0.05	0.01	0.01	0.10	0.33	0.66	0.82	0.29	0.32	0.18	0.24
3-Party		60	60	0.33	0.33	0.00	0.00	0.00	0.33	0.33	0.33	0.30	0.13	0.33	0.50	0.63
4-Party		60	60	0.25	0.50	0.00	0.00	0.00	0.25	0.25	0.45	0.40	0.15	0.25	0.33	0.43
2-Party(1/3)		60	60	0.56	0.33	0.24	0.10	0.11	0.56	0.56	0.15	0.18	0.09	0.76	0.89	0.99
Cyclic		60	60	0.50	0.00	0.00	0.00	0.00	0.25	0.39	0.46	0.57	0.22	0.50	0.33	0.58
Diagonal		60	60	0.02	0.97	0.00	0.00	0.00	0.02	0.02	0.97	0.97	0.62	0.02	0.02	0.01
Triangle		60	60	0.51	0.50	0.49	0.33	0.50	0.50	0.51	0.40	0.32	0.18	0.01	0.14	0.47
N(Tri,0.6)		60	60	0.51	0.21	0.19	0.06	0.15	0.18	0.34	0.90	0.77	0.27	0.00	0.05	0.38
1/2-ID/IC		60	60	0.50	0.51	0.51	0.27	0.27	0.30	0.51	0.40	0.50	0.19	0.09	0.25	0.45

Election	Icon	C	V	satr	av-agr	cntr-agr	pair-agr	pcc-agr	pcc ⁺ -agr	jacc-agr	cntr-div	pcc-div	out-div	cntr-pol	pcc-pol	pair-pol
1/2-IC		60	60	0.50	0.12	0.11	0.02	0.02	0.07	0.34	0.80	0.90	0.29	0.08	0.04	0.18
1/4-IC		60	60	0.25	0.51	0.00	0.02	0.02	0.07	0.15	0.98	0.91	0.30	0.00	0.04	0.15
Lin-IC		60	60	0.50	0.08	0.07	0.01	0.05	0.10	0.30	0.96	0.88	0.27	0.00	0.04	0.37
IC ($p = 0.1$), 1		100	1000	0.10	0.80	0.00	0.00	0.00	0.04	0.05	1.00	0.98	0.39	0.00	0.01	0.08
IC ($p = 0.1$), 2		100	1000	0.10	0.80	0.00	0.00	0.00	0.04	0.05	1.00	0.98	0.39	0.00	0.01	0.08
IC ($p = 0.3$), 1		100	1000	0.30	0.41	0.00	0.00	0.00	0.04	0.18	1.00	0.98	0.33	0.00	0.01	0.10
IC ($p = 0.3$), 2		100	1000	0.30	0.40	0.00	0.00	0.00	0.04	0.18	1.00	0.98	0.33	0.00	0.01	0.10
IC ($p = 0.5$), 1		100	1000	0.50	0.03	0.02	0.00	0.00	0.04	0.34	0.92	0.98	0.32	0.05	0.01	0.10
IC ($p = 0.5$), 2		100	1000	0.50	0.03	0.02	0.00	0.00	0.04	0.34	0.92	0.98	0.32	0.05	0.01	0.10
IC ($p = 0.7$), 1		100	1000	0.70	0.39	0.00	0.00	0.00	0.04	0.54	1.00	0.98	0.33	0.00	0.01	0.10
IC ($p = 0.7$), 2		100	1000	0.70	0.40	0.00	0.00	0.00	0.04	0.54	1.00	0.98	0.33	0.00	0.01	0.10
IC ($p = 0.9$), 1		100	1000	0.90	0.80	0.00	0.00	0.00	0.04	0.82	1.00	0.98	0.39	0.00	0.01	0.08
IC ($p = 0.9$), 2		100	1000	0.90	0.80	0.00	0.00	0.00	0.04	0.82	1.00	0.98	0.39	0.00	0.01	0.08
Lin-IC, 1		100	1000	0.50	0.02	0.02	0.00	0.04	0.08	0.29	0.99	0.94	0.28	0.00	0.01	0.35
Lin-IC, 2		100	1000	0.50	0.02	0.02	0.00	0.04	0.08	0.29	0.99	0.94	0.28	0.00	0.01	0.35
Lin-IC, 3		100	1000	0.50	0.02	0.02	0.00	0.03	0.07	0.29	0.99	0.95	0.28	0.00	0.01	0.35
Lin-IC, 4		100	1000	0.50	0.02	0.02	0.00	0.04	0.08	0.29	0.99	0.94	0.28	0.00	0.01	0.34
Lin-IC, 5		100	1000	0.50	0.02	0.02	0.00	0.04	0.08	0.29	0.99	0.95	0.28	0.00	0.01	0.35
Resampling ($\phi = 0.0045$)		100	1000	0.50	1.00	1.00	0.99	0.99	0.99	0.99	0.00	0.01	0.01	0.00	0.00	0.01
Resampling ($\phi = 0.0095$)		100	1000	0.93	1.00	0.98	0.98	0.98	0.98	1.00	0.02	0.01	0.01	0.00	0.00	0.01
Resampling ($\phi = 0.0155$)		100	1000	0.12	0.99	0.97	0.97	0.97	0.97	0.94	0.03	0.03	0.02	0.00	0.00	0.02
Resampling ($\phi = 0.0761$)		100	1000	0.55	0.93	0.92	0.86	0.86	0.86	0.88	0.08	0.14	0.08	0.00	0.00	0.05
Resampling ($\phi = 0.0954$)		100	1000	0.04	0.99	0.81	0.81	0.82	0.82	0.72	0.16	0.14	0.10	0.00	0.03	0.02
Resampling ($\phi = 0.1139$)		100	1000	0.30	0.90	0.84	0.79	0.79	0.79	0.74	0.16	0.21	0.11	0.00	0.00	0.06
Resampling ($\phi = 0.1307$)		100	1000	0.41	0.87	0.84	0.75	0.76	0.76	0.75	0.16	0.24	0.13	0.00	0.00	0.06
Resampling ($\phi = 0.1444$)		100	1000	0.19	0.91	0.77	0.73	0.74	0.74	0.65	0.23	0.25	0.14	0.00	0.01	0.05

Election	Icon	C	V	satr	av-agr	cntr-agr	pair-agr	pcc-agr	pcc ⁺ -agr	jacc-agr	cntr-div	pcc-div	out-div	cntr-pol	pcc-pol	pair-pol
Resampling ($\phi = 0.1829$)		100	1000	0.64	0.83	0.77	0.67	0.67	0.67	0.79	0.23	0.32	0.17	0.00	0.00	0.07
Resampling ($\phi = 0.2044$)		100	1000	0.34	0.82	0.73	0.63	0.64	0.64	0.61	0.27	0.35	0.18	0.00	0.00	0.07
Resampling ($\phi = 0.2051$)		100	1000	0.48	0.79	0.79	0.63	0.63	0.63	0.68	0.21	0.36	0.18	0.00	0.00	0.08
Resampling ($\phi = 0.2381$)		100	1000	0.72	0.81	0.65	0.58	0.58	0.58	0.79	0.35	0.41	0.20	0.00	0.01	0.07
Resampling ($\phi = 0.2605$)		100	1000	0.67	0.77	0.65	0.55	0.55	0.55	0.74	0.35	0.44	0.21	0.00	0.01	0.08
Resampling ($\phi = 0.3078$)		100	1000	0.41	0.70	0.64	0.48	0.48	0.48	0.53	0.36	0.51	0.23	0.00	0.01	0.09
Resampling ($\phi = 0.3157$)		100	1000	0.22	0.78	0.51	0.47	0.47	0.47	0.42	0.49	0.51	0.24	0.00	0.01	0.07
Resampling ($\phi = 0.3281$)		100	1000	0.99	0.99	0.42	0.58	0.94	0.94	0.99	0.43	0.02	0.12	0.00	0.05	0.01
Resampling ($\phi = 0.3476$)		100	1000	0.19	0.78	0.44	0.42	0.43	0.43	0.37	0.55	0.55	0.25	0.01	0.01	0.07
Resampling ($\phi = 0.3492$)		100	1000	0.58	0.66	0.60	0.43	0.43	0.43	0.61	0.40	0.56	0.24	0.00	0.01	0.09
Resampling ($\phi = 0.3925$)		100	1000	0.59	0.62	0.54	0.37	0.37	0.37	0.59	0.46	0.62	0.25	0.00	0.01	0.10
Resampling ($\phi = 0.4575$)		100	1000	0.32	0.60	0.38	0.29	0.30	0.30	0.35	0.62	0.69	0.27	0.00	0.01	0.09
Resampling ($\phi = 0.4626$)		100	1000	0.57	0.54	0.47	0.29	0.29	0.29	0.53	0.53	0.70	0.27	0.00	0.01	0.10
Resampling ($\phi = 0.4887$)		100	1000	0.86	0.77	0.15	0.27	0.27	0.27	0.82	0.80	0.70	0.30	0.03	0.02	0.07
Resampling ($\phi = 0.4889$)		100	1000	0.47	0.51	0.48	0.26	0.26	0.26	0.44	0.51	0.72	0.27	0.00	0.01	0.10
Resampling ($\phi = 0.5196$)		100	1000	0.87	0.77	0.09	0.23	0.23	0.24	0.82	0.85	0.74	0.31	0.07	0.02	0.07
Resampling ($\phi = 0.5344$)		100	1000	0.44	0.48	0.41	0.22	0.22	0.22	0.39	0.59	0.76	0.28	0.00	0.01	0.10
Resampling ($\phi = 0.5850$)		100	1000	0.03	0.94	0.00	0.17	0.23	0.24	0.12	0.89	0.66	0.43	0.17	0.07	0.04
Resampling ($\phi = 0.6007$)		100	1000	0.60	0.42	0.29	0.16	0.16	0.17	0.49	0.71	0.82	0.29	0.00	0.01	0.10
Resampling ($\phi = 0.6028$)		100	1000	0.58	0.41	0.30	0.16	0.16	0.16	0.48	0.70	0.83	0.29	0.00	0.01	0.10
Resampling ($\phi = 0.6386$)		100	1000	0.66	0.43	0.15	0.13	0.13	0.14	0.55	0.83	0.85	0.30	0.00	0.01	0.10
Resampling ($\phi = 0.6599$)		100	1000	0.37	0.38	0.17	0.12	0.12	0.12	0.29	0.82	0.87	0.30	0.00	0.01	0.10
Resampling ($\phi = 0.6785$)		100	1000	0.93	0.86	0.00	0.11	0.11	0.13	0.88	0.94	0.84	0.40	0.05	0.03	0.06
Resampling ($\phi = 0.6964$)		100	1000	0.97	0.93	0.00	0.11	0.15	0.16	0.94	0.94	0.76	0.45	0.01	0.06	0.04
Resampling ($\phi = 0.7170$)		100	1000	0.20	0.60	0.00	0.08	0.08	0.09	0.15	0.98	0.90	0.33	0.01	0.01	0.09
Resampling ($\phi = 0.7187$)		100	1000	0.89	0.78	0.00	0.08	0.08	0.10	0.82	0.97	0.89	0.36	0.00	0.02	0.08

Election	Icon	C	V	satr	av-agr	cntr-agr	pair-agr	pcc-agr	pcc ⁺ -agr	jacc-agr	cntr-div	pcc-div	out-div	cntr-pol	pcc-pol	pair-pol
Resampling ($\phi = 0.7261$)		100	1000	0.90	0.80	0.00	0.08	0.08	0.10	0.83	0.98	0.89	0.37	0.01	0.02	0.07
Resampling ($\phi = 0.7706$)		100	1000	0.92	0.84	0.00	0.05	0.05	0.08	0.86	0.97	0.91	0.40	0.02	0.02	0.07
Resampling ($\phi = 0.7714$)		100	1000	0.54	0.23	0.16	0.05	0.05	0.07	0.40	0.83	0.93	0.31	0.00	0.01	0.10
Resampling ($\phi = 0.7846$)		100	1000	0.86	0.72	0.00	0.05	0.05	0.07	0.77	0.99	0.92	0.36	0.00	0.02	0.08
Resampling ($\phi = 0.7869$)		100	1000	0.51	0.21	0.19	0.05	0.05	0.07	0.37	0.80	0.94	0.31	0.00	0.01	0.10
Resampling ($\phi = 0.8090$)		100	1000	0.28	0.43	0.00	0.04	0.04	0.06	0.18	0.99	0.94	0.32	0.01	0.01	0.10
Resampling ($\phi = 0.8420$)		100	1000	0.07	0.87	0.00	0.02	0.02	0.06	0.05	0.99	0.93	0.43	0.01	0.03	0.06
Resampling ($\phi = 0.8536$)		100	1000	0.85	0.70	0.00	0.02	0.02	0.05	0.74	1.00	0.95	0.36	0.00	0.01	0.09
Resampling ($\phi = 0.8616$)		100	1000	0.13	0.73	0.00	0.02	0.02	0.05	0.08	1.00	0.96	0.36	0.00	0.01	0.08
Resampling ($\phi = 0.8667$)		100	1000	0.19	0.63	0.00	0.02	0.02	0.05	0.11	1.00	0.96	0.34	0.00	0.01	0.09
Resampling ($\phi = 0.9096$)		100	1000	0.16	0.68	0.00	0.01	0.01	0.05	0.09	1.00	0.97	0.36	0.00	0.01	0.09
Resampling ($\phi = 0.9149$)		100	1000	0.52	0.09	0.06	0.01	0.01	0.05	0.35	0.91	0.97	0.32	0.02	0.01	0.10
Resampling ($\phi = 0.9387$)		100	1000	0.54	0.09	0.01	0.00	0.00	0.04	0.37	0.95	0.98	0.32	0.02	0.01	0.10
Resampling ($\phi = 0.9467$)		100	1000	0.26	0.47	0.00	0.00	0.00	0.04	0.15	1.00	0.98	0.33	0.00	0.01	0.10
Resampling ($\phi = 0.9676$)		100	1000	0.56	0.11	0.00	0.00	0.00	0.04	0.39	0.97	0.98	0.32	0.01	0.01	0.10
Resampling ($\phi = 0.9960$)		100	1000	1.00	1.00	0.00	0.00	0.98	0.98	1.00	1.00	0.02	0.63	0.00	0.00	0.01
Noisy AN ($\phi = 0.0292$)		100	100	0.50	0.02	0.02	0.00	0.00	0.47	0.48	0.22	0.26	0.11	0.95	0.94	0.94
Noisy AN ($\phi = 0.0952$)		100	100	0.50	0.03	0.02	0.00	0.00	0.41	0.44	0.27	0.38	0.16	0.88	0.83	0.83
Noisy AN ($\phi = 0.0988$)		100	100	0.51	0.04	0.03	0.00	0.00	0.41	0.44	0.28	0.39	0.16	0.87	0.81	0.81
Noisy AN ($\phi = 0.1458$)		100	100	0.50	0.04	0.04	0.00	0.00	0.37	0.42	0.30	0.45	0.18	0.82	0.74	0.74
Noisy AN ($\phi = 0.1735$)		100	100	0.50	0.04	0.04	0.00	0.00	0.34	0.41	0.33	0.46	0.19	0.78	0.69	0.69
Noisy AN ($\phi = 0.2060$)		100	100	0.50	0.04	0.04	0.00	0.00	0.32	0.40	0.36	0.49	0.20	0.75	0.63	0.64
Noisy AN ($\phi = 0.2238$)		100	100	0.50	0.05	0.05	0.00	0.00	0.31	0.40	0.36	0.54	0.20	0.74	0.62	0.63
Noisy AN ($\phi = 0.2538$)		100	100	0.50	0.05	0.04	0.00	0.00	0.29	0.38	0.39	0.55	0.21	0.70	0.57	0.57
Noisy AN ($\phi = 0.3319$)		100	100	0.51	0.06	0.05	0.01	0.01	0.23	0.37	0.46	0.64	0.22	0.61	0.45	0.46
Noisy AN ($\phi = 0.3395$)		100	100	0.50	0.06	0.05	0.01	0.01	0.22	0.37	0.47	0.65	0.22	0.60	0.44	0.45

Election	Icon	C	V	satr	av-agr	cntr-agr	pair-agr	pcc-agr	pcc ⁺ -agr	jacc-agr	cntr-div	pcc-div	out-div	cntr-pol	pcc-pol	pair-pol
Noisy AN ($\phi = 0.4155$)		100	100	0.50	0.07	0.07	0.01	0.01	0.18	0.36	0.52	0.71	0.23	0.51	0.34	0.36
Noisy AN ($\phi = 0.4475$)		100	100	0.50	0.06	0.06	0.01	0.01	0.16	0.35	0.55	0.76	0.23	0.49	0.32	0.34
Noisy AN ($\phi = 0.5728$)		100	100	0.51	0.07	0.06	0.01	0.01	0.10	0.35	0.65	0.84	0.24	0.36	0.19	0.23
Noisy AN ($\phi = 0.5966$)		100	100	0.50	0.07	0.06	0.01	0.01	0.09	0.34	0.67	0.84	0.24	0.33	0.17	0.21
Noisy AN ($\phi = 0.6129$)		100	100	0.50	0.08	0.07	0.01	0.01	0.09	0.34	0.66	0.86	0.24	0.32	0.17	0.21
Noisy AN ($\phi = 0.6266$)		100	100	0.50	0.07	0.06	0.01	0.01	0.08	0.35	0.69	0.85	0.25	0.30	0.15	0.20
Noisy AN ($\phi = 0.6957$)		100	100	0.50	0.08	0.08	0.01	0.01	0.07	0.35	0.77	0.89	0.24	0.21	0.10	0.17
Noisy AN ($\phi = 0.7028$)		100	100	0.50	0.07	0.07	0.01	0.01	0.06	0.34	0.74	0.90	0.25	0.20	0.10	0.17
Noisy AN ($\phi = 0.7337$)		100	100	0.49	0.07	0.06	0.01	0.01	0.06	0.33	0.82	0.91	0.25	0.18	0.08	0.16
Noisy AN ($\phi = 0.7346$)		100	100	0.49	0.08	0.06	0.01	0.01	0.06	0.33	0.78	0.91	0.24	0.18	0.08	0.16
Noisy AN ($\phi = 0.7541$)		100	100	0.50	0.09	0.09	0.01	0.01	0.06	0.34	0.80	0.91	0.25	0.15	0.07	0.15
Noisy AN ($\phi = 0.8031$)		100	100	0.50	0.08	0.08	0.01	0.01	0.05	0.34	0.83	0.93	0.25	0.10	0.05	0.15
Noisy AN ($\phi = 0.8212$)		100	100	0.50	0.07	0.07	0.01	0.01	0.05	0.34	0.84	0.94	0.25	0.04	0.04	0.14
Noisy AN ($\phi = 0.9024$)		100	100	0.50	0.07	0.07	0.01	0.01	0.05	0.34	0.86	0.94	0.25	0.06	0.03	0.14
Noisy AN ($\phi = 0.9666$)		100	100	0.49	0.07	0.06	0.01	0.01	0.05	0.33	0.86	0.94	0.25	0.06	0.03	0.14
2-Party (0.09, 0.04)		100	100	0.88	0.92	0.67	0.64	0.85	0.92	0.92	0.07	0.03	0.03	0.33	0.15	0.53
2-Party (0.11, 0.61)		100	100	0.41	0.22	0.06	0.02	0.05	0.52	0.52	0.19	0.19	0.07	0.94	0.95	1.00
2-Party (0.13, 0.11)		100	100	0.79	0.78	0.48	0.41	0.61	0.80	0.80	0.10	0.08	0.05	0.52	0.39	0.79
2-Party (0.14, 0.61)		100	100	0.42	0.22	0.07	0.02	0.05	0.52	0.52	0.19	0.19	0.08	0.93	0.95	1.00
2-Party (0.16, 0.82)		100	100	0.28	0.64	0.36	0.27	0.41	0.70	0.70	0.13	0.12	0.06	0.64	0.59	0.91
2-Party (0.27, 0.58)		100	100	0.46	0.16	0.09	0.02	0.03	0.51	0.51	0.18	0.19	0.07	0.91	0.97	1.00
2-Party (0.28, 0.36)		100	100	0.56	0.28	0.18	0.06	0.08	0.54	0.54	0.16	0.18	0.07	0.82	0.92	1.00
2-Party (0.28, 0.45)		100	100	0.52	0.10	0.06	0.01	0.01	0.50	0.51	0.19	0.20	0.08	0.94	0.99	1.00
2-Party (0.32, 0.46)		100	100	0.51	0.08	0.05	0.01	0.01	0.50	0.50	0.19	0.20	0.08	0.95	0.99	1.00
2-Party (0.43, 0.47)		100	100	0.50	0.06	0.05	0.00	0.00	0.50	0.50	0.19	0.20	0.08	0.95	1.00	1.00
2-Party (0.46, 0.92)		100	100	0.47	0.84	0.83	0.70	0.71	0.85	0.85	0.03	0.06	0.03	0.17	0.29	0.71

Election	Icon	C	V	satr	av-agr	cntr-agr	pair-agr	pcc-agr	pcc+-agr	jacc-agr	cntr-div	pcc-div	out-div	cntr-pol	pcc-pol	pair-pol
2-Party (0.47, 0.34)		100	100	0.51	0.32	0.31	0.10	0.10	0.55	0.55	0.14	0.18	0.07	0.69	0.90	0.99
2-Party (0.52, 0.68)		100	100	0.51	0.36	0.35	0.13	0.13	0.56	0.56	0.13	0.17	0.07	0.65	0.87	0.99
2-Party (0.56, 0.98)		100	100	0.56	0.96	0.95	0.92	0.92	0.96	0.96	0.01	0.02	0.01	0.05	0.08	0.39
2-Party (0.57, 0.16)		100	100	0.46	0.68	0.65	0.46	0.46	0.73	0.73	0.07	0.11	0.05	0.35	0.54	0.89
2-Party (0.58, 0.63)		100	100	0.52	0.26	0.23	0.07	0.07	0.53	0.53	0.15	0.19	0.07	0.77	0.93	1.00
2-Party (0.62, 0.23)		100	100	0.44	0.54	0.47	0.28	0.29	0.65	0.65	0.11	0.14	0.06	0.53	0.71	0.96
2-Party (0.74, 0.28)		100	100	0.39	0.44	0.29	0.16	0.19	0.60	0.60	0.14	0.16	0.07	0.71	0.81	0.98
2-Party (0.79, 0.52)		100	100	0.51	0.04	0.02	0.00	0.00	0.50	0.50	0.20	0.20	0.08	0.98	1.00	1.00
2-Party (0.85, 0.23)		100	100	0.31	0.54	0.26	0.17	0.29	0.65	0.65	0.15	0.14	0.07	0.74	0.71	0.96
2-Party (0.90, 0.47)		100	100	0.48	0.06	0.01	0.00	0.00	0.50	0.50	0.20	0.20	0.08	0.99	1.00	1.00
2-Party (0.91, 0.31)		100	100	0.34	0.38	0.10	0.05	0.14	0.57	0.57	0.18	0.17	0.08	0.90	0.86	0.99
2-Party (0.92, 0.75)		100	100	0.71	0.50	0.14	0.09	0.25	0.62	0.62	0.17	0.15	0.07	0.86	0.75	0.97
2-Party (0.92, 0.19)		100	100	0.24	0.62	0.21	0.16	0.38	0.69	0.69	0.16	0.12	0.06	0.79	0.62	0.92
2-Party (0.95, 0.39)		100	100	0.40	0.22	0.03	0.01	0.05	0.52	0.52	0.19	0.19	0.08	0.97	0.95	1.00
Party-List 1		100	100	0.33	0.34	0.00	0.12	0.10	0.38	0.38	0.26	0.24	0.09	0.76	0.61	0.65
Party-List 2		100	100	0.36	0.50	0.30	0.22	0.20	0.48	0.48	0.19	0.21	0.08	0.42	0.55	0.69
Party-List 3		100	100	0.27	0.67	0.38	0.36	0.33	0.49	0.49	0.21	0.21	0.08	0.31	0.42	0.57
Party-List 4		100	100	0.57	0.36	0.24	0.13	0.19	0.50	0.50	0.19	0.19	0.08	0.58	0.67	0.91
Party-List 5		100	100	0.12	0.77	0.00	0.26	0.26	0.34	0.34	0.39	0.25	0.09	0.56	0.40	0.27
Party-List 6		100	100	0.31	0.44	0.11	0.14	0.17	0.44	0.44	0.22	0.21	0.09	0.18	0.59	0.66
Party-List 7		100	100	0.39	0.22	0.00	0.00	0.03	0.34	0.34	0.29	0.26	0.09	0.53	0.67	0.88
Party-List 8		100	100	0.14	0.72	0.00	0.07	0.08	0.21	0.21	0.58	0.44	0.13	0.18	0.29	0.25
Party-List 9		100	100	0.20	0.71	0.29	0.29	0.36	0.50	0.50	0.25	0.22	0.08	0.38	0.30	0.51
Party-List 10		100	100	0.22	0.56	0.00	0.01	0.06	0.25	0.25	0.48	0.38	0.12	0.17	0.35	0.55
Party-List 11		100	100	0.37	0.27	0.02	0.09	0.08	0.41	0.41	0.24	0.24	0.09	0.77	0.64	0.72
Party-List 12		100	100	0.31	0.38	0.00	0.05	0.08	0.34	0.34	0.32	0.28	0.10	0.45	0.48	0.70

Election	Icon	C	V	satr	av-agr	cntr-agr	pair-agr	pcc-agr	pcc ⁺ -agr	jacc-agr	cntr-div	pcc-div	out-div	cntr-pol	pcc-pol	pair-pol
Party-List 13		100	100	0.17	0.66	0.00	0.16	0.16	0.30	0.30	0.48	0.32	0.11	0.04	0.40	0.34
Party-List 14		100	100	0.24	0.61	0.19	0.20	0.36	0.55	0.55	0.20	0.15	0.08	0.64	0.52	0.65
Party-List 15		100	100	0.22	0.58	0.03	0.12	0.16	0.35	0.35	0.34	0.31	0.10	0.49	0.30	0.61
Party-List 16		100	100	0.19	0.64	0.06	0.09	0.25	0.43	0.43	0.32	0.23	0.09	0.29	0.33	0.69
Party-List 17		100	100	0.35	0.30	0.00	0.02	0.05	0.34	0.34	0.32	0.26	0.10	0.41	0.60	0.73
Party-List 18		100	100	0.30	0.40	0.00	0.03	0.08	0.37	0.37	0.31	0.26	0.10	0.47	0.53	0.69
Party-List 19		100	100	0.28	0.44	0.00	0.03	0.02	0.28	0.28	0.39	0.34	0.11	0.36	0.47	0.53
Party-List 20		100	100	0.27	0.46	0.00	0.03	0.14	0.37	0.37	0.32	0.25	0.10	0.40	0.49	0.65
Party-List 21		100	100	0.14	0.71	0.00	0.11	0.17	0.29	0.29	0.55	0.37	0.12	0.06	0.35	0.36
Party-List 22		100	100	0.22	0.56	0.00	0.11	0.14	0.33	0.33	0.33	0.29	0.10	0.52	0.45	0.55
Party-List 23		100	100	0.29	0.42	0.01	0.05	0.19	0.41	0.41	0.29	0.21	0.10	0.55	0.57	0.74
Party-List 24		100	100	0.32	0.36	0.00	0.13	0.11	0.36	0.36	0.30	0.26	0.09	0.60	0.57	0.62
Party-List 25		100	100	0.27	0.49	0.06	0.08	0.14	0.40	0.40	0.28	0.25	0.09	0.50	0.47	0.75
ID Mixture ($k = 2, p = 0.1$)		100	1000	0.12	0.99	0.94	0.94	0.92	0.92	0.92	0.01	0.02	0.01	0.06	0.08	0.08
ID Mixture ($k = 2, p = 0.3$)		100	1000	0.32	0.58	0.35	0.44	0.44	0.51	0.57	0.13	0.11	0.06	0.65	0.56	0.50
ID Mixture ($k = 2, p = 0.5$)		100	1000	0.50	0.66	0.66	0.57	0.57	0.57	0.70	0.07	0.09	0.05	0.34	0.43	0.47
ID Mixture ($k = 2, p = 0.7$)		100	1000	0.74	0.83	0.66	0.64	0.65	0.66	0.85	0.07	0.07	0.04	0.34	0.35	0.38
ID Mixture ($k = 2, p = 0.9$)		100	1000	0.90	0.92	0.56	0.61	0.69	0.69	0.93	0.09	0.06	0.04	0.44	0.31	0.20
ID Mixture ($k = 3, p = 0.1$)		100	1000	0.10	0.83	0.14	0.46	0.46	0.47	0.49	0.19	0.12	0.06	0.78	0.46	0.18
ID Mixture ($k = 3, p = 0.3$)		100	1000	0.32	0.55	0.30	0.37	0.37	0.39	0.49	0.19	0.17	0.08	0.45	0.42	0.43
ID Mixture ($k = 3, p = 0.5$)		100	1000	0.47	0.78	0.76	0.63	0.63	0.64	0.75	0.06	0.09	0.05	0.18	0.28	0.50
ID Mixture ($k = 3, p = 0.7$)		100	1000	0.71	0.80	0.66	0.62	0.61	0.63	0.83	0.08	0.09	0.05	0.29	0.33	0.41
ID Mixture ($k = 3, p = 0.9$)		100	1000	0.88	0.86	0.45	0.60	0.59	0.59	0.91	0.13	0.10	0.05	0.09	0.33	0.17
ID Mixture ($k = 4, p = 0.1$)		100	1000	0.12	0.77	0.06	0.34	0.32	0.38	0.39	0.27	0.20	0.08	0.56	0.42	0.23
ID Mixture ($k = 4, p = 0.3$)		100	1000	0.32	0.61	0.40	0.41	0.41	0.45	0.53	0.15	0.15	0.07	0.46	0.42	0.46
ID Mixture ($k = 4, p = 0.5$)		100	1000	0.55	0.37	0.30	0.22	0.22	0.26	0.52	0.26	0.30	0.11	0.31	0.30	0.45

Election	Icon	C	V	satr	av-agr	cntr-agr	pair-agr	pcc-agr	pcc ⁺ -agr	jacc-agr	cntr-div	pcc-div	out-div	cntr-pol	pcc-pol	pair-pol
ID Mixture ($k = 4, p = 0.7$)		100	1000	0.68	0.55	0.30	0.33	0.33	0.33	0.67	0.26	0.23	0.09	0.28	0.32	0.38
ID Mixture ($k = 4, p = 0.9$)		100	1000	0.93	0.86	0.02	0.29	0.30	0.33	0.90	0.32	0.22	0.09	0.47	0.38	0.14
ID Mixture ($k = 5, p = 0.1$)		100	1000	0.10	0.80	0.02	0.22	0.21	0.26	0.27	0.43	0.32	0.11	0.26	0.29	0.16
ID Mixture ($k = 5, p = 0.3$)		100	1000	0.32	0.72	0.56	0.51	0.51	0.51	0.59	0.13	0.14	0.07	0.25	0.32	0.42
ID Mixture ($k = 5, p = 0.5$)		100	1000	0.49	0.38	0.36	0.22	0.22	0.29	0.49	0.23	0.29	0.11	0.32	0.37	0.48
ID Mixture ($k = 5, p = 0.7$)		100	1000	0.68	0.52	0.25	0.28	0.28	0.31	0.65	0.27	0.27	0.10	0.13	0.32	0.41
ID Mixture ($k = 5, p = 0.9$)		100	1000	0.91	0.84	0.10	0.37	0.39	0.40	0.90	0.39	0.19	0.09	0.52	0.37	0.16
IAM Mixture ($k = 2$), 1		100	1000	0.53	0.48	0.45	0.30	0.31	0.31	0.51	0.54	0.68	0.26	0.00	0.01	0.11
IAM Mixture ($k = 2$), 2		100	1000	0.49	0.46	0.45	0.26	0.26	0.26	0.46	0.51	0.65	0.27	0.06	0.10	0.16
IAM Mixture ($k = 2$), 3		100	1000	0.50	0.34	0.34	0.18	0.18	0.20	0.43	0.54	0.69	0.28	0.15	0.15	0.19
IAM Mixture ($k = 2$), 4		100	1000	0.50	0.43	0.43	0.26	0.26	0.27	0.46	0.52	0.65	0.27	0.02	0.10	0.17
IAM Mixture ($k = 2$), 5		100	1000	0.50	0.47	0.47	0.29	0.29	0.30	0.48	0.50	0.64	0.26	0.02	0.08	0.18
IAM Mixture ($k = 3$), 1		100	1000	0.52	0.36	0.34	0.18	0.18	0.20	0.44	0.56	0.68	0.28	0.01	0.13	0.21
IAM Mixture ($k = 3$), 2		100	1000	0.51	0.40	0.38	0.22	0.22	0.23	0.45	0.57	0.69	0.27	0.06	0.08	0.18
IAM Mixture ($k = 3$), 3		100	1000	0.44	0.29	0.20	0.11	0.11	0.14	0.34	0.64	0.74	0.29	0.12	0.13	0.18
IAM Mixture ($k = 3$), 4		100	1000	0.50	0.32	0.32	0.14	0.15	0.17	0.40	0.60	0.71	0.28	0.08	0.13	0.20
IAM Mixture ($k = 3$), 5		100	1000	0.48	0.28	0.26	0.12	0.12	0.15	0.38	0.59	0.72	0.29	0.18	0.15	0.20
IAM Mixture ($k = 4$), 1		100	1000	0.49	0.32	0.31	0.14	0.14	0.16	0.40	0.59	0.70	0.28	0.09	0.12	0.19
IAM Mixture ($k = 4$), 2		100	1000	0.55	0.42	0.35	0.23	0.24	0.25	0.49	0.62	0.70	0.27	0.02	0.06	0.17
IAM Mixture ($k = 4$), 3		100	1000	0.48	0.33	0.30	0.15	0.15	0.17	0.39	0.62	0.73	0.29	0.07	0.11	0.19
IAM Mixture ($k = 4$), 4		100	1000	0.51	0.33	0.31	0.16	0.16	0.19	0.43	0.57	0.68	0.28	0.14	0.15	0.23
IAM Mixture ($k = 4$), 5		100	1000	0.52	0.40	0.37	0.21	0.21	0.22	0.46	0.57	0.69	0.28	0.01	0.09	0.18
IAM Mixture ($k = 5$), 1		100	1000	0.49	0.31	0.30	0.14	0.14	0.16	0.40	0.58	0.72	0.29	0.14	0.14	0.19
IAM Mixture ($k = 5$), 2		100	1000	0.47	0.26	0.22	0.10	0.11	0.13	0.36	0.65	0.75	0.29	0.11	0.08	0.18
IAM Mixture ($k = 5$), 3		100	1000	0.48	0.23	0.20	0.08	0.08	0.11	0.35	0.68	0.79	0.30	0.12	0.11	0.17
IAM Mixture ($k = 5$), 4		100	1000	0.52	0.28	0.25	0.11	0.11	0.14	0.41	0.65	0.75	0.29	0.07	0.10	0.18

Election	Icon	C	V	satr	av-agr	cntr-agr	pair-agr	pcc-agr	pcc ⁺ -agr	jacc-agr	cntr-div	pcc-div	out-div	cntr-pol	pcc-pol	pair-pol
IAM Mixture ($k = 5$), 5		100	1000	0.51	0.26	0.24	0.10	0.10	0.13	0.39	0.65	0.76	0.29	0.10	0.11	0.18
2D-Euclidean ($r = 0.117$), 1		100	1000	0.04	0.92	0.00	0.00	0.05	0.09	0.03	0.97	0.88	0.37	0.02	0.04	0.06
2D-Euclidean ($r = 0.117$), 2		100	1000	0.04	0.92	0.00	0.00	0.02	0.05	0.03	0.96	0.91	0.37	0.03	0.04	0.05
2D-Euclidean ($r = 0.117$), 3		100	1000	0.04	0.92	0.00	0.00	0.05	0.08	0.03	0.96	0.87	0.36	0.03	0.05	0.05
2D-Euclidean ($r = 0.117$), 4		100	1000	0.04	0.93	0.00	0.00	0.02	0.06	0.03	0.96	0.90	0.37	0.01	0.04	0.05
2D-Euclidean ($r = 0.117$), 5		100	1000	0.04	0.92	0.00	0.00	0.06	0.09	0.03	0.97	0.88	0.37	0.02	0.03	0.06
2D-Euclidean ($r = 0.117$), 6		100	1000	0.04	0.92	0.00	0.00	0.03	0.07	0.03	0.96	0.90	0.37	0.02	0.04	0.05
2D-Euclidean ($r = 0.117$), 7		100	1000	0.04	0.92	0.00	0.00	0.04	0.08	0.03	0.97	0.89	0.37	0.01	0.04	0.06
2D-Euclidean ($r = 0.117$), 8		100	1000	0.04	0.92	0.00	0.00	0.05	0.09	0.03	0.96	0.87	0.37	0.03	0.05	0.06
2D-Euclidean ($r = 0.117$), 9		100	1000	0.04	0.92	0.00	0.00	0.05	0.09	0.03	0.96	0.87	0.37	0.03	0.04	0.06
2D-Euclidean ($r = 0.117$), 10		100	1000	0.04	0.92	0.00	0.00	0.04	0.07	0.03	0.96	0.89	0.37	0.01	0.04	0.06
2D-Euclidean ($r = 0.167$), 1		100	1000	0.08	0.85	0.00	0.00	0.01	0.07	0.05	0.93	0.86	0.30	0.02	0.07	0.10
2D-Euclidean ($r = 0.167$), 2		100	1000	0.08	0.85	0.00	0.00	0.00	0.07	0.05	0.94	0.87	0.31	0.04	0.07	0.10
2D-Euclidean ($r = 0.167$), 3		100	1000	0.08	0.84	0.00	0.00	0.01	0.08	0.05	0.94	0.86	0.31	0.04	0.07	0.10
2D-Euclidean ($r = 0.167$), 4		100	1000	0.08	0.84	0.00	0.00	0.03	0.09	0.05	0.94	0.84	0.31	0.04	0.07	0.10
2D-Euclidean ($r = 0.167$), 5		100	1000	0.07	0.85	0.00	0.01	0.00	0.07	0.05	0.94	0.87	0.31	0.05	0.07	0.09
2D-Euclidean ($r = 0.167$), 6		100	1000	0.08	0.85	0.00	0.00	0.00	0.07	0.05	0.94	0.87	0.31	0.03	0.07	0.09
2D-Euclidean ($r = 0.167$), 7		100	1000	0.08	0.85	0.00	0.00	0.00	0.07	0.05	0.94	0.87	0.31	0.03	0.07	0.08
2D-Euclidean ($r = 0.167$), 8		100	1000	0.08	0.85	0.00	0.00	0.01	0.07	0.05	0.94	0.86	0.30	0.03	0.07	0.11
2D-Euclidean ($r = 0.167$), 9		100	1000	0.08	0.85	0.00	0.00	0.00	0.07	0.05	0.93	0.86	0.31	0.03	0.07	0.09
2D-Euclidean ($r = 0.167$), 10		100	1000	0.08	0.85	0.00	0.00	0.02	0.09	0.05	0.92	0.84	0.31	0.03	0.07	0.10
2D-Euclidean (Uniform r), 1		100	1000	0.19	0.61	0.00	0.02	0.19	0.28	0.10	0.87	0.66	0.25	0.07	0.10	0.36
2D-Euclidean (Uniform r), 2		100	1000	0.18	0.63	0.00	0.02	0.23	0.31	0.10	0.89	0.62	0.25	0.06	0.10	0.35
2D-Euclidean (Uniform r), 3		100	1000	0.18	0.64	0.00	0.01	0.22	0.30	0.10	0.88	0.62	0.25	0.06	0.10	0.36
2D-Euclidean (Uniform r), 4		100	1000	0.19	0.62	0.00	0.01	0.21	0.30	0.10	0.87	0.62	0.25	0.10	0.12	0.36
2D-Euclidean (Uniform r), 5		100	1000	0.19	0.63	0.00	0.02	0.21	0.29	0.10	0.88	0.63	0.25	0.06	0.10	0.34

Election	Icon	C	V	satr	av-agr	cntr-agr	pair-agr	pcc-agr	pcc ⁺ -agr	jacc-agr	cntr-div	pcc-div	out-div	cntr-pol	pcc-pol	pair-pol
2D-Euclidean (Uniform r), 6		100	1000	0.18	0.63	0.00	0.02	0.21	0.29	0.10	0.88	0.64	0.25	0.08	0.10	0.36
2D-Euclidean (Uniform r), 7		100	1000	0.19	0.62	0.00	0.02	0.24	0.32	0.10	0.87	0.61	0.25	0.07	0.10	0.37
2D-Euclidean (Uniform r), 8		100	1000	0.18	0.64	0.00	0.02	0.22	0.30	0.10	0.89	0.63	0.25	0.07	0.09	0.35
2D-Euclidean (Uniform r), 9		100	1000	0.19	0.63	0.00	0.01	0.21	0.29	0.10	0.88	0.63	0.25	0.07	0.10	0.35
2D-Euclidean (Uniform r), 10		100	1000	0.19	0.61	0.00	0.01	0.19	0.27	0.10	0.86	0.65	0.25	0.07	0.10	0.37
2D-Euclidean (Top10), 1		100	1000	0.10	0.80	0.00	0.01	0.01	0.09	0.08	0.97	0.81	0.27	0.02	0.10	0.09
2D-Euclidean (Top10), 2		100	1000	0.10	0.80	0.00	0.01	0.01	0.09	0.08	0.95	0.81	0.27	0.00	0.10	0.09
2D-Euclidean (Top10), 3		100	1000	0.10	0.80	0.00	0.01	0.01	0.09	0.08	0.96	0.81	0.28	0.00	0.09	0.09
2D-Euclidean (Top10), 4		100	1000	0.10	0.80	0.00	0.01	0.01	0.09	0.08	0.94	0.79	0.27	0.06	0.11	0.09
2D-Euclidean (Top10), 5		100	1000	0.10	0.80	0.00	0.01	0.01	0.09	0.08	0.97	0.82	0.27	0.00	0.09	0.09
2D-Euclidean (Top10), 6		100	1000	0.10	0.80	0.00	0.01	0.01	0.09	0.08	0.96	0.81	0.27	0.00	0.10	0.09
2D-Euclidean (Top10), 7		100	1000	0.10	0.80	0.00	0.01	0.01	0.09	0.08	0.96	0.81	0.27	0.00	0.10	0.09
2D-Euclidean (Top10), 8		100	1000	0.10	0.80	0.00	0.01	0.01	0.09	0.07	0.98	0.83	0.28	0.00	0.09	0.09
2D-Euclidean (Top10), 9		100	1000	0.10	0.80	0.00	0.01	0.01	0.09	0.08	0.97	0.82	0.27	0.00	0.09	0.09
2D-Euclidean (Top10), 10		100	1000	0.10	0.80	0.00	0.01	0.01	0.09	0.08	0.97	0.82	0.27	0.00	0.10	0.09
2D-Euclidean (Uniform Top), 1		100	1000	0.49	0.19	0.17	0.05	0.07	0.17	0.32	0.84	0.73	0.23	0.00	0.16	0.42
2D-Euclidean (Uniform Top), 2		100	1000	0.51	0.20	0.18	0.06	0.08	0.17	0.34	0.84	0.73	0.24	0.00	0.15	0.42
2D-Euclidean (Uniform Top), 3		100	1000	0.48	0.19	0.15	0.05	0.06	0.16	0.30	0.84	0.73	0.24	0.00	0.16	0.42
2D-Euclidean (Uniform Top), 4		100	1000	0.49	0.21	0.20	0.06	0.08	0.17	0.32	0.83	0.72	0.24	0.00	0.15	0.41
2D-Euclidean (Uniform Top), 5		100	1000	0.51	0.21	0.19	0.06	0.08	0.18	0.34	0.82	0.72	0.24	0.00	0.15	0.41
2D-Euclidean (Uniform Top), 6		100	1000	0.50	0.20	0.20	0.06	0.09	0.17	0.33	0.83	0.72	0.24	0.00	0.16	0.41
2D-Euclidean (Uniform Top), 7		100	1000	0.49	0.20	0.18	0.05	0.07	0.17	0.32	0.83	0.73	0.24	0.00	0.16	0.41
2D-Euclidean (Uniform Top), 8		100	1000	0.51	0.22	0.20	0.07	0.10	0.18	0.34	0.83	0.72	0.23	0.00	0.15	0.42
2D-Euclidean (Uniform Top), 9		100	1000	0.51	0.23	0.22	0.07	0.10	0.18	0.34	0.82	0.71	0.23	0.00	0.15	0.41
2D-Euclidean (Uniform Top), 10		100	1000	0.49	0.22	0.21	0.06	0.08	0.18	0.32	0.82	0.72	0.24	0.00	0.16	0.41
Chopin 1		84	17	0.37	0.41	0.22	0.19	0.21	0.22	0.33	0.67	0.64	0.17	0.08	0.08	0.25

Election	Icon	C	V	satr	av-agr	cntr-agr	pair-agr	pcc-agr	pcc ⁺ -agr	jacc-agr	cntr-div	pcc-div	out-div	cntr-pol	pcc-pol	pair-pol
Chopin 2		40	17	0.47	0.28	0.24	0.11	0.12	0.16	0.37	0.70	0.69	0.26	0.00	0.10	0.29
Chopin 3		20	17	0.53	0.36	0.32	0.17	0.18	0.22	0.44	0.55	0.61	0.38	0.00	0.13	0.30
Chopin Final		11	17	0.51	0.25	0.24	0.11	0.36	0.41	0.38	0.71	0.47	0.50	0.00	0.10	0.38
French 2007 All		12	2836	0.18	0.63	0.00	0.15	0.26	0.34	0.21	0.73	0.50	0.55	0.32	0.20	0.27
French 2012 Louvigny		10	930	0.25	0.53	0.07	0.16	0.22	0.31	0.25	0.67	0.54	0.58	0.08	0.21	0.33
French 2012 Saint-Etienne		10	387	0.26	0.50	0.05	0.14	0.20	0.30	0.24	0.67	0.55	0.61	0.26	0.23	0.34
French 2012 Strasbourg		10	1023	0.27	0.51	0.07	0.16	0.28	0.37	0.25	0.72	0.51	0.57	0.07	0.19	0.34
French 2017 All		11	20076	0.24	0.62	0.23	0.28	0.34	0.40	0.32	0.57	0.45	0.46	0.09	0.17	0.30
French 2022 All		12	1379	0.26	0.57	0.18	0.23	0.30	0.37	0.29	0.66	0.51	0.53	0.10	0.15	0.31
Jester 0.25		100	73421	0.25	0.54	0.08	0.17	0.17	0.18	0.24	0.79	0.75	0.29	0.12	0.07	0.14
Jester 0.5		100	73421	0.49	0.51	0.51	0.31	0.31	0.32	0.49	0.42	0.58	0.24	0.05	0.09	0.27
Jester 0.75		100	73421	0.74	0.77	0.56	0.53	0.53	0.55	0.80	0.36	0.35	0.15	0.12	0.15	0.32
MovieLens 4.0		50	423	0.57	0.43	0.34	0.22	0.27	0.27	0.49	0.65	0.69	0.34	0.00	0.04	0.25
MovieLens 4.5		50	423	0.32	0.44	0.14	0.19	0.21	0.23	0.27	0.78	0.74	0.36	0.08	0.03	0.22
MovieLens 5.0		50	423	0.19	0.62	0.00	0.13	0.24	0.26	0.15	0.90	0.70	0.37	0.09	0.04	0.28
Dolphins Network		62	62	0.10	0.80	0.00	0.03	0.02	0.08	0.07	0.93	0.82	0.32	0.02	0.11	0.13
Football Network		115	115	0.10	0.80	0.00	0.00	0.00	0.07	0.07	0.97	0.86	0.26	0.00	0.08	0.08
Karate Club Network		34	34	0.16	0.67	0.01	0.09	0.11	0.18	0.16	0.80	0.64	0.36	0.18	0.20	0.28
Les Miserables Network		77	77	0.10	0.80	0.00	0.07	0.06	0.11	0.09	0.85	0.73	0.26	0.07	0.13	0.19
Polblogs Network		1490	1490	0.01	0.99	0.00	0.02	0.80	0.80	0.30	0.98	0.20	0.32	0.01	0.00	0.04
September11 Network		19	19	0.24	0.53	0.00	0.05	0.04	0.17	0.19	0.74	0.61	0.41	0.13	0.22	0.30
Strike Network		24	24	0.17	0.65	0.00	0.02	0.02	0.14	0.14	0.81	0.66	0.38	0.09	0.19	0.22
Polkadot 18		929	20621	0.01	0.98	0.00	0.06	0.45	0.46	0.09	0.96	0.46	0.26	0.02	0.06	0.02
Polkadot 28		931	22252	0.01	0.98	0.00	0.05	0.54	0.54	0.11	0.97	0.40	0.27	0.05	0.04	0.02
Polkadot 137		993	19442	0.01	0.98	0.00	0.08	0.46	0.47	0.09	0.97	0.47	0.24	0.00	0.05	0.02
Polkadot 158		1002	20603	0.01	0.98	0.00	0.05	0.49	0.50	0.10	0.98	0.45	0.26	0.00	0.04	0.02

Election	Icon	C	V	satr	av-agr	cntr-agr	pair-agr	pcc-agr	pcc ⁺ -agr	jacc-agr	cntr-div	pcc-div	out-div	cntr-pol	pcc-pol	pair-pol
Polkadot 193		1051	23318	0.01	0.99	0.00	0.05	0.54	0.55	0.12	0.97	0.38	0.24	0.04	0.05	0.02
Polkadot 252		1013	30608	0.01	0.98	0.00	0.04	0.52	0.52	0.10	0.97	0.43	0.24	0.01	0.04	0.02
Polkadot 302		995	36726	0.01	0.98	0.00	0.05	0.58	0.59	0.13	0.98	0.38	0.26	0.00	0.03	0.02
Polkadot 353		983	41974	0.01	0.98	0.00	0.05	0.60	0.60	0.14	0.97	0.32	0.28	0.01	0.08	0.02
Polkadot 401		924	46340	0.01	0.98	0.00	0.03	0.62	0.63	0.15	0.98	0.34	0.26	0.00	0.02	0.02
Polkadot 496		973	44918	0.01	0.99	0.00	0.04	0.60	0.61	0.14	0.95	0.35	0.27	0.00	0.03	0.02
Survey Colours		8	140	0.75	0.49	0.00	0.14	0.52	0.58	0.64	0.78	0.35	0.60	0.24	0.09	0.35
Survey Fruits		8	140	0.82	0.64	0.00	0.04	0.67	0.72	0.70	0.86	0.25	0.68	0.08	0.06	0.44
Survey Pets		8	140	0.59	0.24	0.08	0.09	0.42	0.50	0.44	0.75	0.42	0.61	0.07	0.14	0.44
Survey Pizzas		8	140	0.66	0.33	0.02	0.10	0.29	0.38	0.53	0.76	0.54	0.67	0.25	0.13	0.37
Survey Subjects		8	112	0.67	0.42	0.13	0.13	0.25	0.35	0.55	0.67	0.54	0.64	0.17	0.15	0.37
Bidding AAMAS 1		613	201	0.01	0.98	0.00	0.01	0.49	0.49	0.09	0.99	0.48	0.36	0.00	0.01	0.03
Bidding AAMAS 2		442	161	0.01	0.98	0.00	0.01	0.45	0.46	0.08	0.99	0.51	0.35	0.01	0.02	0.03
Bidding AAMAS 3		526	667	0.02	0.96	0.00	0.01	0.25	0.26	0.03	0.99	0.71	0.36	0.01	0.03	0.05
Bidding CS conf 1		54	31	0.10	0.81	0.00	0.09	0.21	0.24	0.11	0.92	0.66	0.31	0.07	0.09	0.14
Bidding CS conf 2		52	24	0.16	0.67	0.00	0.07	0.07	0.12	0.14	0.91	0.76	0.29	0.04	0.11	0.17
Bidding CS conf 3		176	146	0.03	0.94	0.00	0.02	0.18	0.20	0.03	0.99	0.77	0.26	0.01	0.03	0.09
Campsongs 1		78	39	0.29	0.45	0.04	0.10	0.11	0.13	0.22	0.89	0.79	0.23	0.03	0.06	0.24
Campsongs 2		8	39	0.31	0.39	0.01	0.06	0.30	0.40	0.21	0.81	0.50	0.65	0.17	0.15	0.39
Campsongs 3		82	56	0.28	0.47	0.05	0.10	0.11	0.13	0.21	0.90	0.81	0.24	0.03	0.04	0.25
Campsongs 4		10	56	0.36	0.30	0.03	0.05	0.06	0.18	0.24	0.80	0.74	0.65	0.11	0.13	0.34
CTU		23	82	0.30	0.40	0.01	0.08	0.10	0.17	0.20	0.85	0.74	0.44	0.06	0.13	0.36
Eurovision 2		18	18	0.56	0.52	0.46	0.32	0.32	0.35	0.55	0.41	0.49	0.36	0.08	0.12	0.31
Eurovision 4		19	20	0.53	0.51	0.48	0.32	0.32	0.33	0.53	0.42	0.51	0.34	0.06	0.11	0.26
Eurovision 9		18	20	0.56	0.40	0.32	0.21	0.21	0.25	0.50	0.53	0.61	0.40	0.09	0.12	0.29
Eurovision 15		21	22	0.48	0.32	0.29	0.17	0.17	0.21	0.41	0.54	0.64	0.36	0.15	0.12	0.28

Election	Icon	C	V	satr	av-agr	cntr-agr	pair-agr	pcc-agr	pcc ⁺ -agr	jacc-agr	cntr-div	pcc-div	out-div	cntr-pol	pcc-pol	pair-pol
Eurovision 20		24	25	0.42	0.50	0.40	0.31	0.31	0.35	0.45	0.48	0.53	0.31	0.06	0.10	0.30
Eurovision 26		24	24	0.42	0.52	0.42	0.32	0.32	0.34	0.45	0.47	0.51	0.31	0.11	0.11	0.28
Eurovision 32		24	39	0.42	0.48	0.37	0.27	0.27	0.28	0.42	0.53	0.57	0.34	0.11	0.12	0.25
Eurovision 36		24	42	0.42	0.45	0.34	0.26	0.26	0.28	0.42	0.52	0.58	0.36	0.10	0.11	0.25
Eurovision 37		28	42	0.36	0.50	0.30	0.26	0.26	0.27	0.37	0.61	0.61	0.32	0.05	0.09	0.21
Eurovision 39		19	21	0.53	0.53	0.50	0.36	0.36	0.37	0.55	0.37	0.46	0.34	0.11	0.13	0.25
Eurovision 41		25	42	0.40	0.45	0.31	0.23	0.23	0.24	0.38	0.59	0.66	0.37	0.09	0.08	0.23
Eurovision 44		25	39	0.40	0.41	0.26	0.19	0.19	0.21	0.36	0.64	0.67	0.37	0.07	0.09	0.24
Eurovision 48		19	21	0.53	0.31	0.27	0.13	0.13	0.20	0.43	0.54	0.64	0.38	0.13	0.16	0.32
Eurovision 52		18	21	0.56	0.38	0.30	0.19	0.19	0.23	0.49	0.56	0.63	0.39	0.10	0.11	0.30
Eurovision 57		16	19	0.62	0.44	0.25	0.22	0.22	0.26	0.56	0.54	0.55	0.40	0.18	0.14	0.30
Eurovision 63		18	42	0.56	0.51	0.45	0.28	0.28	0.31	0.53	0.49	0.59	0.42	0.06	0.08	0.29
Eurovision 66		18	42	0.56	0.38	0.30	0.20	0.20	0.24	0.49	0.58	0.64	0.44	0.11	0.10	0.27
Eurovision 68		26	86	0.38	0.40	0.22	0.17	0.17	0.21	0.34	0.68	0.70	0.41	0.07	0.11	0.23
Eurovision 69		19	44	0.53	0.42	0.39	0.22	0.22	0.25	0.47	0.52	0.65	0.43	0.07	0.08	0.27
Eurovision 72		17	40	0.53	0.38	0.34	0.19	0.19	0.23	0.46	0.55	0.67	0.44	0.07	0.09	0.28
Posters A		17	65	0.32	0.38	0.04	0.08	0.15	0.22	0.23	0.82	0.70	0.52	0.01	0.10	0.32
Posters B		17	60	0.33	0.41	0.11	0.09	0.16	0.23	0.24	0.80	0.70	0.51	0.05	0.10	0.29
Sushi 10		100	5000	0.10	0.80	0.00	0.09	0.09	0.10	0.10	1.00	0.88	0.35	0.00	0.02	0.05
Sushi ≤ 10		100	5000	0.04	0.92	0.00	0.06	0.06	0.08	0.05	0.97	0.87	0.42	0.03	0.04	0.05
Pol.is 1		54	339	0.07	0.85	0.00	0.08	0.39	0.41	0.09	0.97	0.56	0.40	0.04	0.03	0.19
Pol.is 2		896	2031	0.06	0.87	0.00	0.11	0.42	0.43	0.09	0.96	0.55	0.21	0.04	0.02	0.19
Pol.is 3		1039	1756	0.04	0.92	0.00	0.10	0.44	0.44	0.09	0.95	0.52	0.22	0.07	0.04	0.14
Pol.is 4		522	1116	0.06	0.87	0.00	0.13	0.37	0.38	0.09	0.92	0.56	0.20	0.06	0.03	0.19
Pol.is 5		2138	3142	0.03	0.95	0.00	0.14	0.50	0.50	0.11	0.98	0.46	0.20	0.02	0.03	0.09
Pol.is 6		613	1503	0.04	0.92	0.00	0.12	0.40	0.41	0.09	0.94	0.55	0.21	0.07	0.03	0.12

Election	Icon	C	V	satr	av-agr	cntr-agr	pair-agr	pcc-agr	pcc+-agr	jacc-agr	cntr-div	pcc-div	out-div	cntr-pol	pcc-pol	pair-pol
Pol.is 7		1452	3616	0.03	0.94	0.00	0.12	0.51	0.52	0.11	0.97	0.45	0.21	0.02	0.02	0.10
Pol.is 8		371	126	0.21	0.59	0.00	0.09	0.18	0.20	0.12	0.89	0.74	0.15	0.11	0.05	0.38
Pol.is 9		50	204	0.26	0.57	0.19	0.22	0.28	0.29	0.26	0.73	0.62	0.32	0.04	0.08	0.25
Pol.is 10		174	448	0.09	0.82	0.00	0.08	0.26	0.27	0.06	0.94	0.70	0.20	0.06	0.03	0.27
Pol.is 11		298	1487	0.01	0.97	0.00	0.09	0.59	0.59	0.13	0.98	0.37	0.15	0.01	0.03	0.05
Pol.is 12		39	26	0.39	0.45	0.29	0.25	0.29	0.30	0.36	0.60	0.56	0.25	0.05	0.09	0.29
Pol.is 13		2162	6289	0.03	0.95	0.10	0.35	0.60	0.60	0.22	0.68	0.33	0.10	0.23	0.05	0.04
Pol.is 14		165	381	0.13	0.74	0.00	0.14	0.25	0.26	0.11	0.90	0.68	0.19	0.09	0.06	0.24
Pol.is 15		316	536	0.10	0.81	0.03	0.25	0.37	0.38	0.17	0.79	0.54	0.16	0.16	0.07	0.17
Pol.is 16		80	117	0.27	0.61	0.27	0.36	0.43	0.44	0.32	0.60	0.48	0.20	0.12	0.07	0.21
Pol.is 17		148	334	0.17	0.70	0.13	0.30	0.39	0.39	0.24	0.74	0.53	0.18	0.13	0.07	0.16
Pol.is 18		71	234	0.28	0.53	0.16	0.27	0.37	0.38	0.26	0.66	0.53	0.23	0.21	0.08	0.27
Pol.is 19		297	404	0.13	0.75	0.06	0.25	0.34	0.34	0.18	0.82	0.59	0.17	0.15	0.06	0.16
Pol.is 20		197	1921	0.08	0.83	0.00	0.14	0.28	0.29	0.08	0.91	0.66	0.20	0.09	0.05	0.18
PB Amsterdam 252		85	8520	0.22	0.57	0.00	0.07	0.07	0.10	0.15	0.95	0.87	0.34	0.04	0.04	0.16
PB Amsterdam 309		58	14411	0.18	0.64	0.00	0.08	0.08	0.11	0.14	0.94	0.86	0.43	0.02	0.05	0.14
PB Amsterdam 417		39	9562	0.32	0.40	0.06	0.09	0.09	0.13	0.23	0.84	0.83	0.48	0.05	0.05	0.20
PB Amsterdam 621		60	10081	0.29	0.44	0.04	0.09	0.10	0.13	0.20	0.88	0.84	0.38	0.08	0.05	0.21
PB Amsterdam 631		62	11519	0.24	0.52	0.01	0.10	0.09	0.12	0.18	0.92	0.85	0.39	0.07	0.05	0.17
PB Budapest 2022 Jozsefvaros		33	520	0.12	0.77	0.00	0.11	0.11	0.16	0.14	0.92	0.77	0.51	0.10	0.09	0.11
PB Budapest 2023 Jozsefvaros		40	918	0.12	0.77	0.05	0.13	0.12	0.16	0.14	0.91	0.80	0.51	0.02	0.05	0.09
PB Budapest 2024 Jozsefvaros		40	864	0.12	0.76	0.00	0.08	0.08	0.13	0.12	0.93	0.84	0.53	0.03	0.06	0.09
PB Budapest 2024		175	28943	0.05	0.90	0.00	0.07	0.07	0.08	0.06	1.00	0.84	0.27	0.00	0.09	0.08
PB Budapest 2025 Ujbuda		40	2040	0.07	0.85	0.00	0.06	0.06	0.11	0.08	0.94	0.81	0.54	0.01	0.09	0.09
PB Lodz 2020		151	51472	0.03	0.95	0.00	0.05	0.04	0.06	0.04	0.95	0.87	0.36	0.00	0.05	0.03
PB Lodz 2022		160	90494	0.02	0.96	0.00	0.04	0.04	0.06	0.04	0.88	0.82	0.34	0.00	0.11	0.03

Election	Icon	C	V	satr	av-agr	cntr-agr	pair-agr	pcc-agr	pcc+-agr	jacc-agr	cntr-div	pcc-div	out-div	cntr-pol	pcc-pol	pair-pol
PB Lodz 2022 Teoflow Wielkopolska		24	16771	0.13	0.88	0.56	0.56	0.55	0.59	0.56	0.37	0.28	0.22	0.05	0.18	0.22
PB Lodz 2023		139	70747	0.03	0.95	0.00	0.04	0.04	0.06	0.04	0.96	0.86	0.39	0.04	0.06	0.03
PB Lodz 2024		163	66045	0.02	0.95	0.00	0.04	0.04	0.05	0.04	0.94	0.87	0.38	0.00	0.05	0.03
PB Toulouse 2019		30	1494	0.09	0.81	0.00	0.06	0.07	0.13	0.11	0.90	0.72	0.56	0.01	0.16	0.14
PB Toulouse 2022 5		10	423	0.18	0.64	0.00	0.09	0.10	0.21	0.18	0.79	0.62	0.64	0.16	0.20	0.26
PB Toulouse 2022 12		10	659	0.18	0.74	0.28	0.32	0.35	0.41	0.35	0.55	0.42	0.47	0.08	0.18	0.24
PB Toulouse 2022		199	4532	0.01	0.97	0.00	0.02	0.02	0.03	0.02	0.97	0.91	0.40	0.01	0.05	0.01
PB Toulouse 2024		183	7260	0.02	0.97	0.00	0.01	0.01	0.03	0.02	1.00	0.94	0.40	0.01	0.02	0.01
PB Warszawa 2021		106	95899	0.08	0.84	0.00	0.11	0.10	0.12	0.10	0.95	0.84	0.36	0.03	0.04	0.07
PB Warszawa 2022		129	81234	0.06	0.87	0.00	0.08	0.07	0.09	0.07	0.97	0.87	0.34	0.00	0.04	0.06
PB Warszawa 2023		138	78904	0.06	0.88	0.00	0.09	0.09	0.10	0.08	0.96	0.85	0.32	0.01	0.04	0.06
PB Warszawa 2024		118	75255	0.07	0.85	0.00	0.10	0.09	0.11	0.09	0.97	0.85	0.34	0.02	0.04	0.06
PB Warszawa 2025		86	67615	0.10	0.80	0.00	0.08	0.08	0.10	0.09	0.97	0.85	0.38	0.00	0.04	0.08
PB Warszawa 2026 Bemowo		35	4714	0.13	0.74	0.00	0.09	0.09	0.14	0.13	0.92	0.81	0.55	0.02	0.07	0.11
PB Warszawa 2026 Bialoleka		30	6542	0.15	0.70	0.00	0.10	0.10	0.15	0.14	0.91	0.81	0.57	0.04	0.06	0.13
PB Warszawa 2026 Bielany		27	4656	0.17	0.66	0.00	0.07	0.07	0.13	0.14	0.90	0.82	0.61	0.05	0.07	0.14
PB Warszawa 2026 Mokotow		17	6902	0.27	0.48	0.03	0.08	0.08	0.16	0.21	0.83	0.79	0.68	0.07	0.08	0.22
PB Warszawa 2026 Ochota		23	3060	0.20	0.59	0.00	0.10	0.10	0.15	0.17	0.86	0.79	0.62	0.10	0.08	0.17
PB Warszawa 2026 Praga Polnoc		25	1746	0.19	0.63	0.01	0.12	0.12	0.16	0.18	0.87	0.78	0.58	0.11	0.07	0.15
PB Warszawa 2026 Praga Poludnie		21	6202	0.21	0.58	0.00	0.08	0.08	0.15	0.17	0.85	0.78	0.64	0.07	0.09	0.19
PB Warszawa 2026 Rembertow		40	1555	0.11	0.78	0.00	0.05	0.05	0.11	0.09	0.93	0.82	0.54	0.06	0.08	0.10
PB Warszawa 2026 Srodmiestcie		10	2113	0.41	0.29	0.13	0.09	0.10	0.20	0.30	0.70	0.73	0.71	0.16	0.13	0.32
PB Warszawa 2026 Targowek		24	3793	0.19	0.62	0.00	0.06	0.06	0.12	0.14	0.90	0.83	0.63	0.04	0.08	0.16
PB Warszawa 2026 Ursus		15	3415	0.30	0.49	0.15	0.18	0.18	0.23	0.28	0.70	0.70	0.60	0.11	0.08	0.25
PB Warszawa 2026 Ursynow		36	5782	0.13	0.74	0.00	0.09	0.09	0.14	0.13	0.91	0.81	0.53	0.08	0.07	0.11
PB Warszawa 2026 Wawer		19	3259	0.22	0.56	0.00	0.11	0.11	0.17	0.19	0.88	0.77	0.64	0.13	0.08	0.20

Election	Icon	C	V	satr	av-agr	cntr-agr	pair-agr	pcc-agr	pcc ⁺ -agr	jacc-agr	cntr-div	pcc-div	out-div	cntr-pol	pcc-pol	pair-pol
PB Warszawa 2026 Wesola		49	1256	0.10	0.81	0.00	0.05	0.05	0.10	0.08	0.92	0.84	0.50	0.03	0.10	0.08
PB Warszawa 2026 Wilanow		8	1367	0.51	0.24	0.23	0.08	0.09	0.21	0.38	0.69	0.74	0.75	0.09	0.13	0.36
PB Warszawa 2026 Wlochy		10	2102	0.44	0.24	0.13	0.08	0.08	0.19	0.33	0.68	0.75	0.74	0.21	0.14	0.33
PB Warszawa 2026 Wola		15	4210	0.30	0.46	0.10	0.14	0.14	0.21	0.26	0.75	0.73	0.64	0.18	0.09	0.25
PB Warszawa 2026 Zoliborz		20	2282	0.23	0.57	0.04	0.15	0.15	0.21	0.22	0.78	0.72	0.56	0.11	0.10	0.21

Table 4: Values of our indices for elections in our main dataset.