




VoteKit: A Python package for computational social choice research

Jennifer Wang

DOI: [10.xxxxxx/draft](https://doi.org/10.xxxxxx/draft)

Software

- [Review](#) 
- [Repository](#) 
- [Archive](#) 

Editor: [Open Journals](#) 

Reviewers:

- [@openjournals](#)

Submitted: 01 January 1970

Published: unpublished

License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC BY 4.0](#)).

Summary

The scholarly study of elections, known as *social choice theory*, centers on the provable properties of voting rules. The practical study of democracy reform focuses on designing or selecting systems of election to produce electoral outcomes that promote legitimacy and broad-based representation. For instance, the dominant electoral system in the United States is a one-person-one-vote/winner-take-all system, sometimes known as PSMD (plurality in single member districts), but there is considerable reform momentum in favor of ranked choice voting because it is thought to mitigate the effects of vote-splitting and strengthen prospects for minority representation.¹ Across the world, systems of election—and prospects for system change—vary substantially. From both a scholarly and a practical perspective, many questions arise about comparing the properties and tendencies of diverse systems of election in a rigorous manner.

VoteKit ([MGGG Redistricting Lab, n.d.-b](#)) is a Python package designed to facilitate just that kind of analysis, bringing together multiple types of functionality. Users can:

1. Create synthetic *preference profiles* (collections of ballots) with a choice of generative models and behavioral parameters;
2. Read in real-world *cast vote records* (CVRs) as observed examples of preference profiles; clean and process ballots, including by deduplication and handling of undervotes and overvotes;
3. Run a variety of *voting rules* to ingest preference profiles and output winner sets and rankings; and
4. Produce a wide range of *summary statistics* and *data visualizations* to compare and analyze profiles and election outcomes.

Statement of need

Social choice theory grew out of welfare economics in the mid-twentieth century and has been recognized as a deep and highly applicable area of economic theory, forming part of the basis for at least four Nobel Prize awards.² Since the 1990s, a new fusion of economics and computer science has emerged under the name of *computational social choice*, studying questions of complexity and design and further advancing the axiomatic study of elections.³ But most of these innovations have been highly abstract, and there has been a significant gap

¹Recent ranked-choice voting reforms include the adoption of instant runoff voting (IRV) in Maine, Alaska, New York City, and single transferable vote (STV) in Portland, Oregon. Advocacy groups claiming various pro-democratic properties of ranked choice include [Campaign Legal Center](#), [FairVote](#), and many others.

²Nobel Laureates with significant work in social choice include Arrow, Sen, Maskin, and Myerson.

³For example, a very active research direction in computational social choice theory has been the development of fairness axioms for approval elections, such as the definition called JR (justified representation) and its relatives, which have been extended to rankings. See ([Aziz et al., 2017](#); [Skowron et al., 2017](#)) and their references.

34 in the literature—and in the landscape of software—between the theory and the practice of
35 democracy.

36 On the software side, researchers have built a multitude of different packages for generating and
37 analyzing elections, and users have had to invest substantial work in cleaning CVRs to make
38 them usable across multiple packages.⁴ VoteKit is built to provide an end-to-end pipeline.

39 Area of need: Generative models

40 For one concrete example of a literature and software gap, consider the construction of
41 *generative models*. This term is often associated with large language models as paradigms
42 of artificial intelligence; here, what is being generated is realistic voting rather than realistic
43 language. In this setting, a generative model of voting is a probability distribution on the
44 set of all possible ballots that can be cast in a given election style; profiles can be sampled
45 from a generative model to produce simulated or synthetic elections. Having sources of rich,
46 varied, and realistic data is essential to an empirically grounded research program to probe
47 the properties of voting rules. Good generative models are also essential to advise reformers
48 deciding between options in a new locality, as they enable generation of synthetic profiles keyed
49 to the scale, demographics, and election specs of that specific place. But most of the models
50 in the literature, like the Impartial Culture model (all permutations of candidates are equally
51 likely) or the Impartial Anonymous Culture model (sampling proportional to volume measure
52 on the simplex of weighted averages of permutations) are mathematically tractable but highly
53 unrealistic. This is bluntly described by Tideman and Plassman in a survey of generative
54 methods: in their words, “None of the 11 models discussed so far are based on the belief that
55 the associated distributions [...] might actually describe rankings in actual elections” (Tideman
56 & Plassmann, 2010). They therefore recommend *spatial models* instead, which themselves are
57 of dubious realism for the selection of political candidates.⁵

58 VoteKit implements many of the models described in those surveys, as well as newer mathe-
59 matical models that give users the ability to generate profiles that are designed to comport
60 with real-world ranking behavior and particularly to generate polarized elections. Two leading
61 choices are based on classic statistical ranking mechanisms, called the Plackett–Luce (PL) and
62 Bradley–Terry (BT) models; another model called the Cambridge Sampler (CS) draws from
63 historical ranking data in Cambridge, MA city council elections. These models have flexible
64 parameters, allowing users to vary voting bloc proportions, candidate strength within slates,
65 and polarization between blocs. These parameters can be specified or randomly sampled.

66 Area of need: Comparison and communication

67 In the realm of democracy reform, groups of stakeholders often ask researchers to provide
68 modeling studies to decide on what shift to make in electoral systems, as the project list below
69 makes clear. VoteKit implements voting rules that stakeholders often seek to compare, with
70 parameters designed to be tailored by the user to the specific locality under study. Available
71 voting rules include:

⁴See for instance the extensive array of open-source tools on the Computational Social Choice (COMSOC) community page (Ulle Endriss and Simon Rey, n.d.) including the widely used collection of ranked data called PrefLib (Ulle Endriss and Simon Rey, n.d.) See also the materials provided by FairVote, including their DataVerse and GitHub (FairVote, n.d.). The ArXiv preprint (Boehmer et al., n.d.) provides an impressively comprehensive list of numerical experiments on elections. The PRAGMA Project (<https://perma.cc/2P6V-8ZER>) echoes our statement of need, noting that the current literature and software falls short in practical applicability and that the understanding of real and synthetic data is “very limited.”

⁵Spatial models assume voters rank by proximity in a metric space defined by issue positions or other attributes; the metric space may be latent, or unknown to voters, but it is presumed to universally govern the way voters rank candidates. See for instance (Burden, 1997), which introduces probabilistic voting keyed to proximity. Though spatial models have been argued to perform adequately to model roll call voting in Congress, their efficacy for selecting political representation is debatable. In a meta-analysis of 163 papers (Boehmer et al., n.d.), the authors report that Impartial Culture and Euclidean (spatial) models make up more than 75% of the election experiments found in 163 papers.

- 72 ■ **Ranking-based (ordinal).** Plurality/SNTV, STV and IRV, (generalized) Borda, Alaska⁶,
73 Top-Two, Dominating sets/Smith method, Condo-Borda⁷, Sequential RCV.
 - 74 ■ **Score-based (cardinal).** Range voting, Cumulative, Limited.
 - 75 ■ **Approval-based (set).** Approval voting, Bloc plurality.
- 76 See generally (Amorós et al., 2016; Emerson, 2013; McCune et al., 2023; Reynolds et al.,
77 2008; Tideman, 1995) for references.
- 78 Reform advocates also need to describe voting mechanisms and their likely outcomes effectively
79 to members of their communities. VoteKit includes a variety of metrics and visualizations
80 intended to facilitate this.

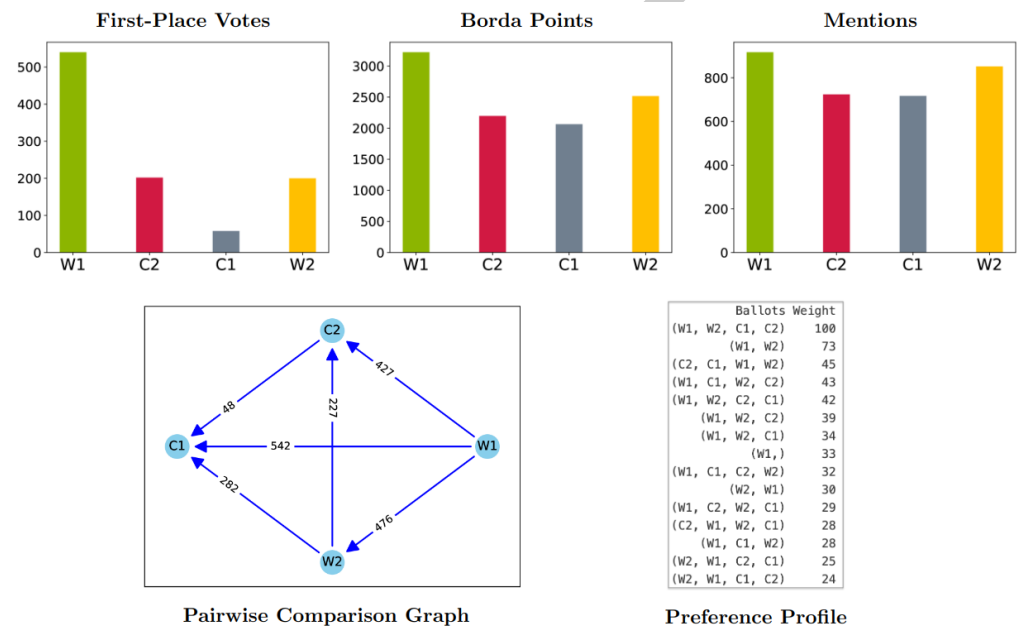


Figure 1: Four visualizations for the same synthetic four-candidate preference profile with 1000 voters. The profile was generated with the CS model, which produces incomplete ballots at a realistic rate.

81 For a given preference profile, basic VoteKit functions provide statistics and visuals for first-
82 place votes, Borda count, and mention frequency, as well as head-to-head comparisons (see
83 Figure 1). The pairwise comparison graph shows head-to-head margins between candidates.
84 For instance, 282 more voters ranked $W_2 > C_1$ than $C_1 > W_2$ in the preference profile used
85 to make the figure. Note that W_1 , who has the most first-place votes by far, is also preferred
86 head-to-head over all alternatives, making them the *Condorcet candidate* in this election.

87 Area of need: Resources for research

88 Previous research works such as (Elkind et al., 2017) have compared properties of earlier
89 generative models; VoteKit facilitates robust comparisons across a more comprehensive and
90 up-to-date list of alternatives. It also offers new analytical tools that will support research
91 on elections. Some examples of more sophisticated functionality are shown in Figure 2. At

⁶Our model of the Alaska method is an SNTV/STV hybrid that uses single non-transferable vote to choose a set of finalists, then runs STV on the same preference profile to fill the seats. Alaska's elections run this with four finalists and one seat; the top-two system runs this with two finalists and one seat.

⁷This system orders candidates within dominating sets by Borda score. Note that this is distinct from Black's method (Black, 2012), which uses Borda score as a backup system in case the smallest dominating set is not a singleton.

left is a *ballot graph*, where nodes are ballots weighted by their frequency in the profile; a recent research paper shows that ballot graphs can be metrized to realize classical statistical ranking distances, like Kendall tau and the Spearman footrule (Duchin & Tapp, 2024). VoteKit also implements a class of election distances, as surveyed in (Boehmer et al., 2022). Choices for measuring the difference between two profiles on the same set of candidates include L^p distance and Wasserstein (earth-mover) distance. At right is a multidimensional scaling (MDS) plot of a different set of data, showing mutual L^1 differences between generated profiles across various selections of model (shown in colors) and candidate strength parameters (shown with symbols), enabling comparisons in the style of (Szufa et al., 2020).

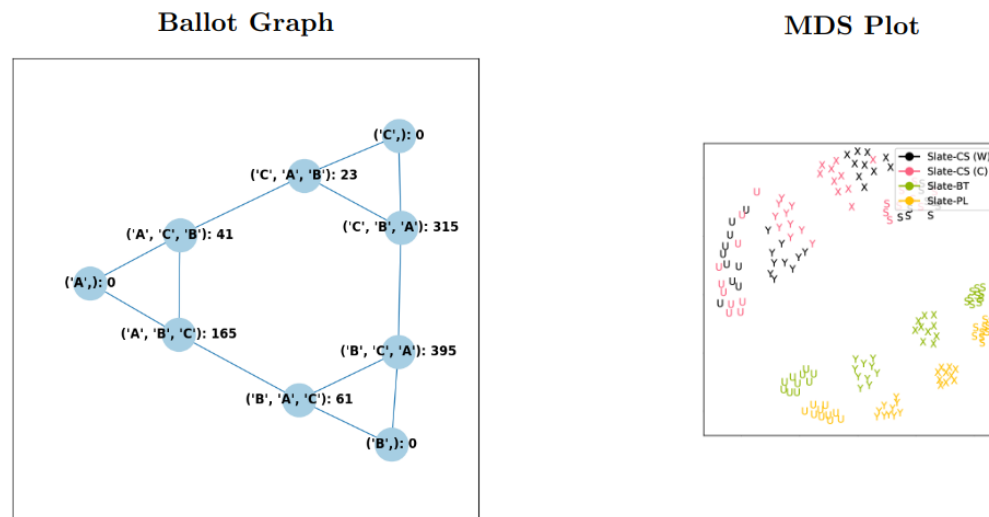


Figure 2: At left, the ballot graph for a 3-candidate election. There is one node per possible ballot, and the weights show the number of instances of that ballot in the profile. At right, a multidimensional scaling (MDS) plot for 160 synthetic profiles made with various generative models and candidate strength parameters for two slates of 3 candidates each. The MDS plot is a low-distortion planar embedding of those 160 profiles and their pairwise differences.

Finally, VoteKit interacts seamlessly with a wide range of actual vote data, such as thousands of political elections collected by FairVote and a cleaned repository of over 1000 Scottish STV local government elections (FairVote, n.d.; MGGG Redistricting Lab, n.d.-a). Previously, the use of real data in election research was often extremely limited; for instance, a recent survey reports that the single most popular “real-life” dataset has been a survey of 5000 respondents’ sushi preferences (Boehmer et al., n.d.).

Projects

A significant number of white papers and scholarly articles have used VoteKit (and its predecessor codebase) in recent years. These include the following.

- A large number of case studies in ranked-choice modeling, such as studies for the city councils of Chicago, IL (MGGG Redistricting Lab, 2019b) and Lowell, MA (MGGG Redistricting Lab, 2019a); the state legislatures of Oregon and Washington (MGGG Redistricting Lab, 2021a, 2021d), and a range of county commissions and school boards across the Pacific Northwest (MGGG Redistricting Lab, 2021c, 2021b);
- A study modeling the impact of proposed legislation called the Fair Representation Act, which would convert U.S. Congressional elections to the single transferable vote system (MGGG Redistricting Lab, 2022);

- 118 ▪ A detailed study isolating the impacts of varying hypotheses about voter behavior and
119 candidate availability on the Massachusetts legislature ([MGGG Redistricting Lab, 2024](#));
- 120 ▪ A peer-reviewed article for an election law audience on the impact of STV elections on
121 minority representation ([Benadè et al., 2021](#));
- 122 ▪ A peer-reviewed article for a CS/econ audience that probes whether STV delivers
123 proportional representation ([Benadè et al., 2024](#)); and
- 124 ▪ A peer-reviewed article for an CS/operations research audience on optimizing to “learn”
125 blocs and slates in real-world elections ([Duchin & Tapp, 2024](#)).

126 Acknowledgements

127 This work was initiated in a research cluster in Summer 2023, funded by the Democracy Fund
128 and graciously hosted at the Faculty of Computing and Data Sciences at Boston University
129 and the Tisch College of Civic Life at Tufts University. Major contributors to the initiation of
130 the project include Brenda Macias, Emarie De La Nuez, Greg Kehne, Jordan Phan, Rory Erlich,
131 James Turk, and David McCune. Earlier code contributions were made by Chanel Richardson,
132 Anthony Pizzimenti, Gabe Schoenbach, Dylan Phelan, Thomas Weighill, Dara Gold, and Amy
133 Becker. The authors also thank Deb Otis, Peter Rock, Jeanne Clelland, and Michael Parsons
134 for helpful feedback. FairVote’s data repository in Dataverse ([https://dataverse.harvard.edu/
135 dataverse/rcv_cvrs](https://dataverse.harvard.edu/dataverse/rcv_cvrs)) and RCV Cruncher code on GitHub ([https://github.com/fairvotereform/
136 rcv_cruncher/](https://github.com/fairvotereform/rcv_cruncher/)) are excellent open-source efforts that were inspirational for the current project.

137 References

- 138 Amorós, P., Puy, M. S., & Martínez, R. (2016). Closed primaries versus top-two primaries.
139 *Public Choice*, 167, 21–35. <https://doi.org/10.1007/s11127-016-0328-5>
- 140 Aziz, H., Brill, M., Conitzer, V., Elkind, E., Freeman, R., & Walsh, T. (2017). Justified
141 representation in approval-based committee voting. *Social Choice and Welfare*, 48(2),
142 461–485.
- 143 Benadè, G., Buck, R., Duchin, M., Gold, D., & Weighill, T. (2021). *Ranked choice voting and
144 proportional representation*.
- 145 Benadè, G., Donnay, C., Duchin, M., & Weighill, T. (2024). Proportionality for ranked voting,
146 in theory and practice. In *Preprint*. <https://mggg.org/PRVTP>.
- 147 Black, D. (2012). *The theory of committees and elections*. Springer Dordrecht.
- 148 Boehmer, N., Faliszewski, P., Janeczko, Ł., Kaczmarczyk, A., Lisowski, G., Pierczyński, G.,
149 Rey, S., Stolicki, D., Szufa, S., & Wąs, T. (n.d.). *Guide to numerical experiments on
150 elections in computational social choice*.
- 151 Boehmer, N., Faliszewski, P., Niedermeier, R., Szufa, S., & Wąs, T. (2022). Understanding
152 distance measures among elections. In L. D. Raedt (Ed.), *Proceedings of the thirty-
153 first international joint conference on artificial intelligence, IJCAI-22* (pp. 102–108).
154 International Joint Conferences on Artificial Intelligence Organization. [https://doi.org/10.
155 24963/ijcai.2022/15](https://doi.org/10.24963/ijcai.2022/15)
- 156 Burden, B. C. (1997). Deterministic and probabilistic voting models. *American Journal of
157 Political Science*, 41(4), 1150–1169. <http://www.jstor.org/stable/2960485>
- 158 Duchin, & Tapp. (2024). Learning blocs and slates from observed elections. *Preprint*.
- 159 Elkind, E., Faliszewski, P., Laslier, J.-F., Skowron, P., Slinko, A., & Talmon, N. (2017). What
160 do multiwinner voting rules do? An experiment over the two-dimensional euclidean domain.

- 161 *Proceedings of the AAAI Conference on Artificial Intelligence*, 31.
- 162 Emerson, P. (2013). The original borda count and partial voting. *Social Choice and Welfare*,
163 40, 353–358. <https://doi.org/10.1007/s00355-011-0603-9>
- 164 FairVote. (n.d.). *RCV cruncher*.
- 165 McCune, D., Martin, E., Latina, G., & Simms, K. (2023). *A comparison of sequential*
166 *ranked-choice voting and single transferable vote*. <https://arxiv.org/abs/2306.17341>
- 167 MGGG Redistricting Lab. (n.d.-a). *Scottish STV election repo*.
- 168 MGGG Redistricting Lab. (n.d.-b). *VoteKit*.
- 169 MGGG Redistricting Lab. (2019a). Findings on the city of lowell’s election systems. In *White*
170 *Paper*. <https://mggg.org/Lowell-Detailed-Report>.
- 171 MGGG Redistricting Lab. (2019b). Study of reform proposals for chicago city council. In
172 *White Paper*. <https://mggg.org/publications/Chicago.pdf>.
- 173 MGGG Redistricting Lab. (2021a). Analysis of election systems for oregon state. In *White*
174 *Paper*. <https://mggg.org/Oregon>.
- 175 MGGG Redistricting Lab. (2021b). Analysis of election systems for the chelan county,
176 washington board of county commissioners. In *White Paper*. [https://mggg.org/Chelan_](https://mggg.org/Chelan_County)
177 [County](https://mggg.org/Chelan_County).
- 178 MGGG Redistricting Lab. (2021c). Analysis of election systems for the tukwila, WA school
179 district. In *White Paper*. <https://mggg.org/Tukwila>.
- 180 MGGG Redistricting Lab. (2021d). Analysis of election systems for washington state. In *White*
181 *Paper*. <https://mggg.org/Washington>.
- 182 MGGG Redistricting Lab. (2022). Modeling the Fair Representation Act. In *White Paper*.
183 <https://mggg.org/FRA-Report>.
- 184 MGGG Redistricting Lab. (2024). Comparing electoral systems for the massachusetts legislature.
185 In *White Paper*. <https://mggg.org/MA-RCV>.
- 186 Reynolds, A., Reilly, B., & Ellis, A. (2008). *Electoral system design: The new international*
187 *IDEA handbook*. International Institute for Democracy; Electoral Assistance.
- 188 Skowron, P., Lackner, M., Brill, M., Peters, D., & Elkind, E. (2017). Proportional rankings.
189 *International Joint Conference on Artificial Intelligence (IJCAI 2017)*.
- 190 Szufa, S., Faliszewski, P., Skowron, P., Slinko, A., & Talmon, N. (2020). Drawing a map of elec-
191 tions in the space of statistical cultures. *Proceedings of the 19th International Conference*
192 *on Autonomous Agents and MultiAgent Systems*, 1341–1349. ISBN: 9781450375184
- 193 Tideman, N. (1995). The single transferable vote. *Journal of Economic Perspectives*, 9(1),
194 27–38. <https://doi.org/10.1257/jep.9.1.27>
- 195 Tideman, N., & Plassmann, F. (2010). *The structure of the election-generating universe*.
196 <https://api.semanticscholar.org/CorpusID:119079594>
- 197 Ulle Endriss and Simon Rey. (n.d.). *COMSOC community site*.