Voting Biases in Decentralized Autonomous Organization (DAO) Governance

Stefano Balietti[1], Pietro Saggese[2,3], and Markus Strohmaier[3,4,5]

[1] Private Sector
[2] IMT School for Advanced Studies Lucca
[3] Complexity Science Hub Vienna
[4] University of Mannheim
[5] GESIS - Leibniz Institute for the Social Sciences

Abstract. Decentralized Autonomous Organizations (DAOs) use token-weighted voting to allocate resources, set protocol rules, and legitimate collective decisions. Yet, support in DAO voting is strikingly concentrated. What happens inside the ballot that produces this concentration? We study DAOs' governance at the proposal-choice level, linking each choice's voting-power share to three observable features: whether it expresses an approval-oriented stance, where it appears in the choice list, and whether it is selected by the proposal author. We find that (i) author-selected choices show the strongest and most robust association with voting-power share, with a 58.8% increase relative to non-author choices; (ii) approval-oriented choices retain a positive but slightly less consistent advantage (27.1%); and (iii) first-listed choices also attract systematically higher shares, consistent with position and order effects (7.7%). Results are robust across several specifications, which include subtracting an author's own voting power from computations. We use *bias* descriptively, to denote systematic associations rather than proven causal distortion. The results shift attention from proposal outcomes alone to the interface and social signals through which choices are presented. In DAO governance, ordering, author signals, and vote visibility should be treated as institutional design choices, not neutral implementation details.

JEL Codes: D71, D72, G23, G34, O33

Keywords: Decentralized Autonomous Organization, DAO, Governance, Bias, Blockchain, Voting, Collective decision-making, Decentralized Finance, DeFi

1 Introduction

Decentralized Autonomous Organizations (DAOs) are a novel blockchain-based governance model that promises to deliver efficient and democratic coordination mechanisms for small and large groups of individuals [66,32]. By employing smart contracts and so-called governance tokens deployed on distributed ledger technologies (DLTs), they automate governance rules and distribute voting rights to community members, boasting a flat hierarchy with no central authority and an open, transparent, and democratic governance model. DAOs test a central promise of digital governance: that transparency and open participation can produce robust collective decisions. Yet public voting alone does not show whether outcomes emerge from substantive participation or from the social cues, presentation choices, and human processes through which proposals are encountered and supported.

DAOs govern a wide range of projects, from decentralized applications such as DeFi protocols [4], to social and funding initiatives run by online communities [22,67], and collaborative projects in virtual economies, including metaverse, play-to-earn, and NFT ecosystems [30,6]. Notably, many of these organizations – especially those controlling DeFi protocols like MakerDAO [43,58], Uniswap [2], Sushiswap [59], and Compound [39] – manage treasuries in the order of tens or hundreds of million of dollars.

Decision-making is executed through voting on so-called *governance proposals* that can affect any aspect of the organization: its technical infrastructure [62], the economic incentives and design [24,20], and the allocation of funds [61,63]—for instance, for hiring new developers or starting a marketing campaign. Governance users can participate in the voting by possessing *governance tokens*, which determine their proportional share of decision-making power (in short, voting power). Prior research has documented several frictions in the DAO decision-making processes. The ownership of governance tokens is highly concentrated [34,49,10,26], and participation in voting is usually low [9], with individuals who possess the potential power to alter outcomes rarely exercising it [29,28]. Prominent users play an influential role in the decision-making process: large voters and vested users (project owners, administrators, developers) vote strategically and influence decision-making processes [53,54,36], while authors of governance proposals engage in insider trading [21].

These frictions coexist with another puzzle: when DAO proposals reach formal voting, support is often highly skewed toward approval or toward a single winning side. Faqir-Rhazoui et al. [27] report high approved-proposal rates across DAOstack, Aragon, and DAOhaus, ranging from around 75% in DAOstack to around 90% in DAOhaus, and argue that off-chain discussion, sponsorship, and the costs of failed proposals may screen out proposals before a vote. In protocol governance, Messias et al. [47] find that Compound proposals receive an average of 88.63% of votes in favor, with a median of 97.66%. In Snapshot governance, Wang et al. [65] show that single-choice voting dominates, binary proposals are the most common pattern, and over 60% of proposals end with one-sided results.

Prior work therefore establishes that DAO voting often records strong support; however, it leaves open which proposal-choice features are associated with that support once voters encounter a ballot. Established literature from related fields, such as online voting and recommender systems, suggests that collective decisions can be shaped by how options are presented, by the social signals attached to them, and by the processes through which proposals reach a vote. First, earlier-listed choices may benefit from interface and order effects, consistent with survey response-order and ballot-order research [38,37,46] – a pattern we refer to as *position bias*. Second, approval-oriented choices may be advantaged by proposal screening and by approve or reject governance structures in which proposals are discussed, sponsored, or revised before formal voting [27,65,47], i.e., *approval bias*. Third, choices selected by proposal authors may reflect author voting power, an informative author signal, or social influence from visible early voting and reputational cues [48,52,25], a pattern we refer to as *author bias*. To our knowledge, however, these forms of bias have not yet been conceptualized and empirically examined in DAO voting systems.

In this paper, we investigate the extent to which the high approval rates and one-sided outcomes observed in Snapshot, a leading governance platform for DAO voting, can be explained by the three biases identified above, and how much these overlap with one another. We note that we use *bias* to denote a systematic association with voting-power share, not as proof that any feature causally distorts preferences. Accordingly, we ask:

RQ1 (Position). Do earlier-listed choices receive a disproportionate share of voting power?

RQ2 (Approval). Are approval-oriented choices systematically favored, and is this advantage distinct from list position?

RQ3 (Author). Do author-selected choices receive a voting-power advantage beyond what position and stance already explain?

We find that author-selected choices gain 58.8 percentage points of voting-power share on average, more than double the 27.1% advantage of approving choices and nearly eight times the 7.7% advantage of first-listed choices. This paper contributes thus large-scale evidence on Snapshot governance at the proposal-choice level. It also introduces a novel stance classification procedure triangulated with external model and human evaluation checks, and it empirically decomposes the associations between voting-power share and author selection, list position, and approval-oriented stance. Notably, our findings are valid across several alternative specifications; even after recomputing estimations by removing author's own voting power, our headline results do not change substantially.

Understanding these associations is crucial to shedding light on the quality and robustness of collective decision-making in DAOs. Systematic advantages for particular kinds of choices may penalize dissenting positions and mask limited deliberation, suggesting that seemingly technical features of the voting process can have institutional consequences. Measuring biases also helps clarify the role of prominent actors, such as proposal authors, whose participation is disproportionately

aligned with support. More broadly, identifying systematic biases is essential to evaluate the legitimacy of DAO governance and, ultimately, to guide improvements in the design of decentralized decision-making systems.

The paper is structured as follows. Section 2 provides background information and reviews the related literature. Section 3 describes the data and outlines the data processing and cleaning procedures. Sections 4 and 5 present the descriptive results and the main empirical analyses. Section 6 concludes the paper discussing the findings and their implications.

2 Background and Literature Review

2.1 Decentralized Autonomous Organizations (DAOs)

Decentralized Autonomous Organizations are a novel organizational structure, designed to offer a decentralized form of governance alternative to traditional corporate structures [32,53]. The key components of a DAO are deployed on distributed ledger technology (DLT): governance rules are encoded in smart contracts that determine who has the power to make the management decisions and establish how the decision makers exercise their power; on-chain governance tokens represent and distribute this decision-making power across community members; and treasuries, also managed via smart contracts, hold the DAO's funds. By relying on blockchain technologies, DAOs aim thus to remove central authorities and support more democratic and participatory governance [7].

Decision making in DAOs. DAO governance decision-making follows a process through which community members collectively propose, evaluate, and implement decisions. This process begins with proposal submission by members (pre-voting period), followed by token-based voting and finally the implementation of the proposed changes (post-voting period).

Pre-voting phase. In this period, a *governance* or *improvement* proposal is created. It submits for collective consideration a proposed change, action, or decision on the DAO structure or the underlying project. The proposal is composed of a body and title, defining the question and the suggested change, and a set of choices, defining the alternative options to decide on.

Proposals can have a direct effect on the organization, its governance, and treasury, or serve a broader scope. The former aim to implement smart contract changes, give a stipend to a member, elect a new board member, or modify protocol parameters (e.g. set a trading fee on a on a decentralized exchange, modify the interest rates in a lending protocol, determine what tokens to list or de-list). The latter are more heterogeneous. They range from opinion polls, e.g. asking the user base about expected future cryptoasset price changes or their interest in certain features of the organization, to selections of options from a number of pre-selected alternatives, e.g. to decide the official logo of the organization, or to select one NFT project from a list of alternatives.

Choices can be expressed as simple binary responses, numbers, strings of text, or other arbitrary content. We define the intended meaning or position expressed by a voting choice within a proposal

as the *stance* of that choice. In most proposals, choices simply indicate whether to approve or reject the proposal content; however, they can also encode more complex stances.

Voting and post-voting phase Once the proposal is created, the voting period begins. Community members who hold governance tokens can vote, and their voting power is proportional to the amount of governance tokens they control. DAOs adopt two alternative approaches to enable voting: on-chain voting records votes directly on the blockchain, and can automatically execute governance decisions through smart contracts; off-chain voting takes place on external platforms like Snapshot [56]. In the latter, votes are authenticated via on-chain addresses and voting power is calculated based on on-chain token balances, but votes are recorded off-chain. While on-chain voting is more secure and transparent, it incurs in high transaction costs, and most DAOs rely on off-chain voting.

Platforms like Snapshot use several voting systems, differing in how voters express preferences and how votes are aggregated (see Table 2 for a description of all voting systems available in Snapshot and their prevalence in our sample). The most widely adopted system is single-choice [65], where users indicate one single preference. More complex voting systems allow users to express preferences for several choices with a single vote. Voters can e.g. select multiple options and assign weights reflecting their relative preference, or rank choices in order of preference.

Finally, in the post-voting period, accepted proposals are executed and implemented to reflect the accepted changes.

2.2 Biases and User Behavior in Online Collaborative Platforms

Online collaborative platforms routinely translate heterogeneous individual judgements into aggregate rankings, ratings, or voting outcomes. This makes them vulnerable to several well-documented behavioral and interface effects. For our purposes, the most relevant channels are: (i) presentation effects, especially the tendency to favor earlier options; (ii) social influence in sequential decision-making; (iii) reputation or authority signals attached to users; and (iv) positive-outcome skew arising from both selection processes before formal voting begins and social-desirability pressures during voting.

Presentation effects In search and recommendation settings, users are more likely to inspect and select items displayed higher in a ranked list [35,23]. Similar effects have been documented in Reddit and Hacker News [57], StackExchange [14,16,41], digital library recommender systems [19], crowdsourcing systems [15], and cultural markets [33,1]. The voting literature provides an even closer analogy: survey response-order effects and ballot-order effects show that the first-listed response or candidate can receive an advantage, especially when voters have limited information or low incentives to deliberate [38,37,46]. In online political primaries, for instance, candidates listed at the top of the screen received a measurable advantage [45]. These findings directly motivate our focus on the ordering of choices in Snapshot proposals. Other interface-induced biases, such

as attention bias, novelty bias, trust bias, and anchoring bias, have also been studied in online settings [23,3,69]; however, Snapshot presents voters with a comparatively simple list of choices, making position bias the most salient interface channel in our setting. Recent work further shows that large language model recommendations can exhibit similar position effects [64,12]; we treat this as adjacent evidence that ordered-choice representations can induce systematic ranking effects, while our behavioral analysis focuses on human voting decisions.

Social influence. Classical models of sequential decision-making show how individuals may rationally follow earlier actors rather than rely on their private information, producing herding or informational cascades [8,11]. Online experiments and observational studies show similar dynamics in digital environments: social influence can increase inequality and unpredictability in cultural markets [55], prior ratings can causally affect later ratings [48], and social signals can shape item popularity alongside content and position [33]. Sequential voting systems can also create herding effects [18], and even limited social influence can undermine the wisdom-of-crowds effect [42]. Because Snapshot votes and intermediate results are often publicly visible during the voting period, DAO voting is a natural environment for such mechanisms. Recent DAO studies are consistent with this view: late voters are more likely to support leading alternatives [68], large voters can strategically influence outcomes [53], and emerging work explicitly studies information cascades in DAO voting [13].

Reputation signals. Reputation and authority provide a third channel. Online platforms often attach visible status signals to users, and reputation systems can affect trust and perceived quality [52]. Prior work has measured authority in social media and Q&A platforms [50,51], linked user reputation to perceived contribution quality [60,40], and shown that reputation can affect persuasion [44]. The closest study to ours, Dev et al. [25], jointly estimates reputation, social influence, and position biases in StackExchange using an instrumental-variable approach. In DAO governance, analogous status signals are attached not only to highly visible voters and delegates but also to proposal authors. An author's vote may therefore be informative, persuasive, or focal, especially when it occurs early and remains visible to later voters.

Positive-outcome skew. Finally, high approval rates may reflect a broader positive-outcome skew rather than voter preferences alone. One channel is social desirability: voters may be more inclined to support approval-oriented choices when approval is perceived as cooperative, constructive, or aligned with community goals [31]. Another channel operates before the vote. DAO governance proposals are often discussed, revised, or informally tested in forums before they reach a formal vote. Empirical studies of DAO platforms report high positive-vote or approval rates across DAOstack, DAOhaus, and Aragon, and explicitly link these patterns to off-chain discussion and the costs of submitting proposals that are likely to fail [27]. Similarly, studies of Snapshot and protocol governance find strong skew toward agreement or in-favor votes [65,47]. In this paper, we therefore use "approval bias" to denote the systematic advantage of approval-oriented choices among observed

Table 1: Raw and cleaned Snapshot dataset, October 2020–November 2023.

	Raw	**Cleaned**	**Removed**	**Reduction**
Spaces	32 995	19 201	13 794	41.81%
Proposals	223 760	127 281	96 479	43.12%
Choices	653 913	398 218	255 695	39.1%
Votes	57 626 386	50 759 796	6 866 590	11.92%

proposals, while recognizing that this advantage may reflect both social-desirability pressures during voting and selection processes before the formal vote.

Taken together, the literature establishes that online collective evaluations are shaped by choice order, visible prior behavior, and user-level authority signals. Our contribution is to bring these channels into the DAO governance setting and quantify their associations at the proposal-choice level. Specifically, we ask how much of the high support observed in Snapshot governance is associated with the stance of a choice, its position in the list, and whether it is selected by the proposal author.

3 Data

The original dataset, spanning from October 2020 to November 2023, was obtained from Snapshot.org [56]. To validate it and ensure consistency, we perform several sanity checks, as discussed in Appendix B. The cleaned dataset contains ∼51 million votes cast on ∼127 thousand proposals. Table 1 reports summary statistics.

3.1 Data Processing

Before proceeding with the analysis, we homogenize two sources of heterogeneity across Snapshot proposals: the voting systems used to cast votes, and the formats used to express choices. Further details for each of the following subsections are reported in the Appendix B.

Multi-preference vote expansion. As Table 2 shows, most proposals (93%) use single-preference voting, where one vote maps to one choice; the remainder use multi-preference systems (e.g., weighted or quadratic voting), where a single vote can be split across several choices. To make these different voting systems comparable, we map multi-preference ballots to choice-level rows using two alternative schemes: *favorite-choice*, which assigns each ballot's unit mass (i.e., the vote) to its top-weighted choice, and *support-fraction*, which divides mass proportionally across all supported choices. Both approaches conserve each ballot's total mass and results are highly comparable. In the remainder of the paper, we rely primarily on the favorite-choice scheme (see Appendix B.2 for details).

Table 2: Snapshot voting systems and their prevalence in the analytic sample

Voting system	Voting system description	Proposals N (%)	Choices N (%)
Single-preference voting systems			
Basic	Pick one from three choices: Yes, No, Abstain.	8975 (7.05%)	26 925 (6.76%)
Single-choice	Pick one from any number of choices.	109 398 (85.95%)	266 893 (67.02%)
Subtotal		118 373 (93%)	293 818 (73.78%)
Multi-preference voting systems			
Weighted	Express both preference and intensity (summing to 100) for any number of choices.	4671 (3.67%)	79 043 (19.85%)
Quadratic	Express both preference and intensity for any number of choices. Voting power grows linearly, while intensity grows quadratically.	1793 (1.41%)	10 591 (2.66%)
Ranked-choice	Rank choices in order of preference.	907 (0.71%)	5031 (1.26%)
Approval	Approve any number of choices with equal weight.	1537 (1.21%)	9735 (2.44%)
Subtotal		8908 (7%)	104 400 (26.22%)

Notes: Percentages are computed over the 127 281 proposals and 398 218 choices retained in cleaned dataset. Voting systems are grouped into single-preference (one vote per ballot) and multi-preference (one ballot spread over several choices).

Choice stance detection. As discussed in Section 2, choices may express similar stances using different formats across proposals, ranging from simple binary responses to numbers, strings of text, and other arbitrary content. For instance, an approving choice can be formulated as a simple 'Yes' string text, a 'thumbs up' emoji, or a longer text answer. We thus devise an approach to classify each choice into a simplified and comparable set of stances: *Approve*, *Reject*, *Abstain*, or *Other*. We implement this categorization approach using three different methods: (i) heuristics, our main tool in the analysis, (ii) large language models, and (iii) crowdworker evaluations. We use the latter two as a validation tool.

Heuristics. The pipeline has three stages. First, a keyword classifier assigns an initial stance to each choice by matching its text against a curated multi-language keyword list. Second, we manually review cases the classifier cannot reliably resolve (such as negated phrasing, abstention language, and choices represented through emojis) and adjust their labels accordingly. Third, we apply a small set of documented corrections at the level of the whole DAO space or the whole proposal to fix systematic mislabeling that the first two stages miss. Finally, we enforce a rule that a governance proposal must contain at least one Approve and one Reject choice for either label to apply: proposals with only Approve (1,945) or only Reject (4,274) choices are relabeled as Other (see Appendix C.1 for details).

Large language models (LLMs) evaluations. To assess the robustness of the heuristic classification, we validate it with external LLM evaluations of both open source and commercial models (Llama

3.1/3.3, GPT-4o, ChatGPT-4o). We instruct each LLM to categorize all proposals and related stances following the prompts reported in Appendix C.2. Notably, this allows us to also examine whether the body of a proposal implicitly or explicitly contains a cue pointing readers toward a voting direction. We treat this *suggested stance* as an LLM-coded textual cue in the proposal text, not as verified author intent or evidence that voters were persuaded. We thus obtain two variables for each LLM that respectively capture whether and in which direction the text suggests a vote, and classify each choice as 'Abstain', 'Approve', 'Other', or 'Reject'.

To measure the agreement across the two classification methods, we compute Cohen's Kappa between our heuristics and each LLM, (using model-specific denominators so Llama 3.3's sparser coverage is handled consistently). The values are 0.69 for Llama 3.1, 0.68 for Llama 3.3, 0.36 for OAI, and 0.61 for ChatGPT, indicating substantial, substantial, fair, and substantial agreement, respectively. Llama 3.1 has the highest agreement with the heuristic-based approach (see Appendix C.2 for details).

Human evaluations. We further validate the heuristic approach by measuring the extent to which it agrees with external human evaluations; this evaluation is implemented with NodeGame [5].

First, we create a stratified multilingual sample of 590 representative random proposals by computing the embeddings of each proposal, reducing their dimensionality using UMAP, and clustering them with the K-means algorithm. Next, using the online labor market for research Prolific, we assign a group of evaluators (N = 375) the task of manually labeling five proposals, so that each proposal is reviewed at least 3 times. Finally, we adopt a majority rule approach to aggregate the evaluations provided by different crowdworkers and assign each stance a unique label ('Abstain', 'Approve', 'Other', or 'Reject'). In case of tie, we leave the label unassigned, and we fill in unassigned stances when the label can be inferred from the remaining choices of the same proposal.

We compare the heuristic-based classification and the crowdworkers' evaluations: the Cohen's Kappa is 0.71, which indicates substantial agreement. Further information on crowdworkers' demographics and the aggregation of their evaluations is reported in Appendix C.3.

4 Biases in Decentralized Governance

Voting power is the variable that determines the outcome of a proposal, namely, the choice with the highest voting power is implemented. Operationally, it is a weighted sum of the individual votes, and the exact formulation varies across DAOs, which are free to experiment and adapt how it is computed. Most DAOs adopt a strategy that assigns voting power to a vote proportionally to the amount of governance tokens held by a given voter. That means that, in principle, a single contrarian vote from somebody with a large weight could significantly shift the distribution of voting power across the choices of a proposal. However, as shown in Table 14, this is rarely the case: voting power and individual votes are highly correlated (95%); hence, in the following, we only focus on voting power, but the results apply to both.

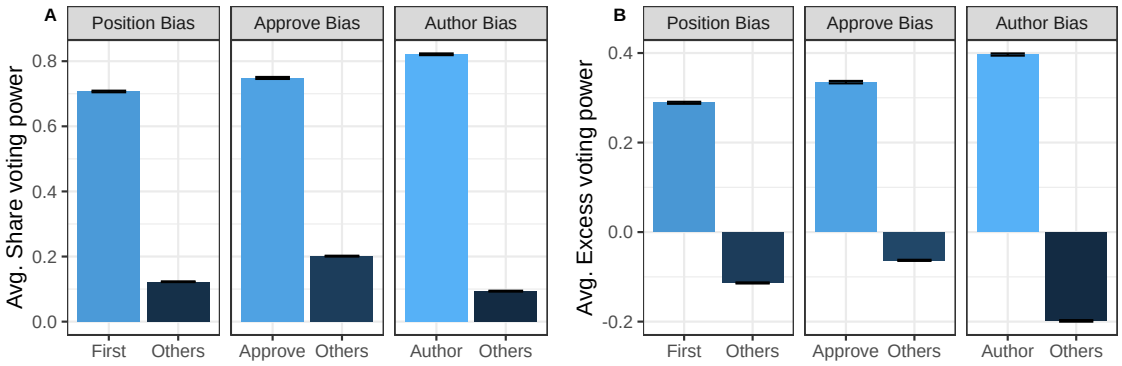


Fig. 1: **Share and excess of choice-level voting power for the three indicators under study**. **A.** Outcome: mean voting-power share for choices that are first-listed, approval-oriented, or selected by the proposal author. **B.** Outcome: mean excess voting power relative to a uniform-within-proposal benchmark. Error bars show 95% confidence intervals for mean estimates. Across both measures, the choice associated with each bias captures far more voting power than the alternative ones, indicating that each channel is associated with a substantial, non-random concentration of support.

Measures. To compare proposal outcomes regardless of how voting power is computed, in all our analyses we focus on two normalized measures computed at the choice level:

- *share*: the voting power of a choice divided by the total amount of voting power of all choices in a proposal; e.g., in a proposal with two choices with voting power 10 and 20, the total voting power is 30, and the share of the two choices is respectively 1/3 and 2/3.
- *excess / deficit*: the difference between the actual share and a uniformly random share; e.g., in a proposal with four choices, the random share is 0.25, the choice A with a share of 0.6 would have an excess voting power of $0.6 - 0.25 = 0.35$, while the choice B with a share of 0.1 would have a deficit of -0.15.

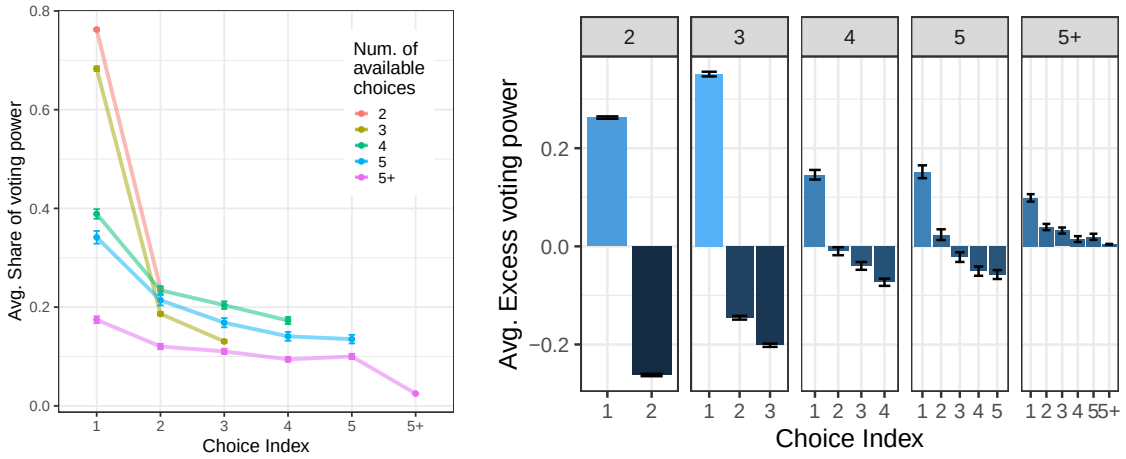
Fig. 1 illustrates the relationship between voting-power share and excess/deficit for the three biases under study in all governance proposals. In this figure, as well as in the following, error bars represent 95% confidence intervals of the mean estimates. Share does not sum to one and excess does not sum to zero in this figure because proposals with different numbers of choices are pooled together; both quantities sum to their respective totals (one and zero) only within proposals sharing the same number of choices, a comparison we hold fixed in the sections that follow.

4.1 Position Bias

We use position bias descriptively for the association between list order and voting-power share. We find a highly imbalanced distribution of voting power across choices, with the first choice seizing

a disproportionate share of voting power across all proposals: 70% across all proposals (see Table 14) with a peak of 79% for proposals of type Basic (see Table 15). This pattern is consistent with strong order effects.

Fig. 2A shows how the share of voting power changes across proposals with different numbers of choices. The association is strongest on proposals with two and three choices; in the latter, the first choice captures around 80% of total voting power. These proposals often follow an approve/reject pattern, as in the Basic voting system, which always has exactly three choices. Fig. 2B shows the excess voting power for proposals with different numbers of choices: the extra voting share captured by the first choice is drawn proportionally from choices with higher indexes. To quantify this pattern, however, we also need to account for stance and author selection, as we do in the next sections.



(a) Voting-power share.

(b) Excess voting power.

Fig. 2: **Choice position and voting power by number of choices in a proposal.** Mean voting-power share (**A**) and mean excess voting power (**B**) by choice position and number of choices in the proposal. Excess is relative to a uniform-within-proposal benchmark by choice position and number of choices. Error bars show 95% confidence intervals for mean estimates. The positional advantage is higher in two- and three-choice proposals, but it can be observed across all ballot lengths.

4.2 Approval Bias

The stance of a choice—i.e., whether it is approving or rejecting the proposal—may influence its voting share. In fact, choices with an Approve stance overwhelmingly capture the largest share of voting power across all proposals (about 80%, see Fig. S10A). However, this dominance is not evenly distributed across all choice positions, but predominantly driven by a large excess of voting power

in the first position (see Fig. 3). All stances get boosted for being placed in the first position, but their excess voting power is much more modest. Overall, at every position the Approve stance has a positive delta compared to other stances, and the same pattern can be observed disaggregating proposals by number of choices (see Fig. S11).

To study the interaction between position and approval-oriented stance, we need to look at how often the Approve stance is placed in position one versus other positions. Fig. 10B and Table 16 show that Approve choices are much more likely to occupy the first choice for proposals with two or three choices, which are more often associated with approve/reject proposals. There are at least two explanations for this observation. First, by design, the vastly popular Basic voting system places Approve choices in position one. Second, proposal authors may place approval-oriented choices first. To better understand this point, we next examine author selection.

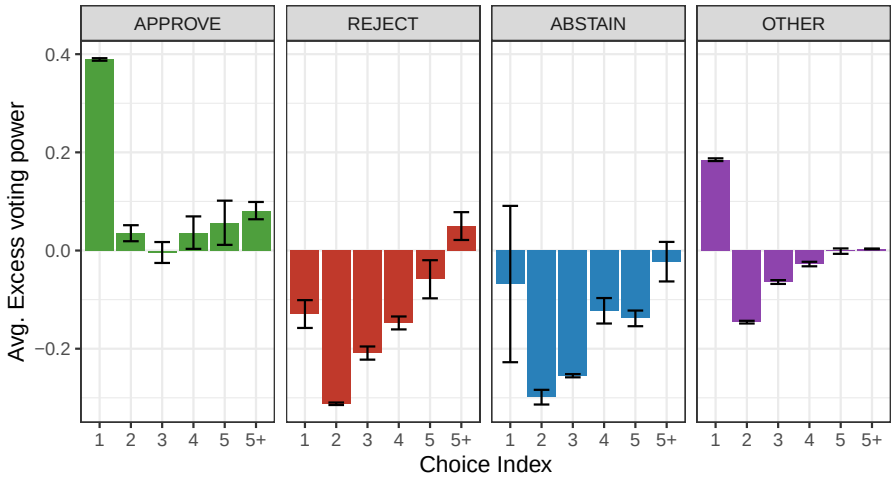


Fig. 3: **Choice stance, list position, and excess voting power.** Excess voting power is plotted separately for each stance and by choice position. Excess is relative to a uniform-within-proposal benchmark. Error bars show 95% confidence intervals for mean estimates. All stances obtain a boost in voting power when placed in position 1, but the Approve stance obtains the largest gain and has limited deficit at other positions.

4.3 Author Bias

Proposal authors might influence outcomes with their own choices. Fig. 4A shows that the choice selected by the author receives the largest amount of voting power for any choice position and for any number of choices. This boost is larger for the first choice than for other choices, suggesting that the author-choice association is larger than the position association, which still plays an important role.

Fig. 4B compares approval stance and author selection, highlighting two descriptive patterns. First, even when disaggregated by stance, author-selected choices show positive excess voting power across positions, although the size varies. Second, the author boost is stable across all choice positions only for the Approve stance, while it is decreasing for Reject and Other stances (see also Fig. S12); interestingly, the highest boost overall is for the Reject choices at position two and three.

Overall, the results in this section suggest a nested descriptive pattern: author-selected choices have the largest voting-power advantage, being at the first position also generates gains, and the effect of approval-oriented choices is partly intertwined with author selection and first placement. In the next section, we evaluate this pattern with an econometric framework.

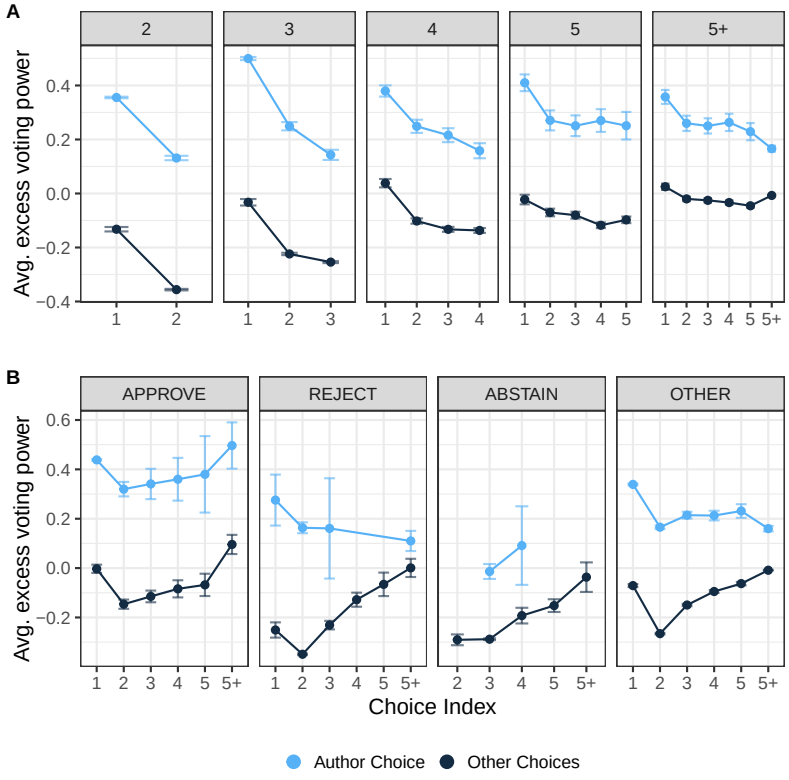


Fig. 4: **Author-selected choices and excess voting power. A.** Excess voting power by number of choices, choice position, and whether the proposal author selected the choice. **B.** Excess voting power by stance, choice position, and whether the proposal author selected the choice; data points averaging 10 observations or less are removed from plot. Excess is relative to a uniform-within-proposal benchmark. Error bars are 95% confidence intervals of the means. In all cases, author-selected choices receive larger voting power compared to each alternative outcome.

Table 3: Overlap among position, approval, and author-choice indicators

	Voting power		Approve		Author choice		All choices	
Choice rank	Mean	SD	Mean	SD	Mean	SD	Share (%)	N
Panel A: all proposals								
1	0.70	0.39	0.40	0.49	0.73	0.44	31.96	127,281
2	0.22	0.35	0.01	0.12	0.21	0.41	31.96	127,281
3	0.14	0.26	0.02	0.15	0.16	0.37	9.87	39,303
4	0.14	0.24	0.03	0.16	0.20	0.40	3.48	13,839
5	0.11	0.22	0.02	0.14	0.18	0.38	2.04	8,130
6	0.09	0.18	0.01	0.11	0.16	0.37	1.28	5,110
7	0.07	0.15	0.02	0.13	0.18	0.39	0.97	3,853
8	0.06	0.14	0.01	0.10	0.18	0.39	0.81	3,209
9	0.05	0.13	0.01	0.11	0.17	0.38	0.69	2,766
10	0.05	0.12	0.01	0.10	0.17	0.37	0.62	2,482
10+	0.01	0.06	0.01	0.09	0.09	0.29	16.31	64,964
Panel B: proposals with up to three choices								
1	0.74	0.37	0.44	0.50	0.77	0.42	44.95	113,442
2	0.23	0.35	0.01	0.09	0.21	0.41	44.95	113,442
3	0.13	0.27	0.01	0.08	0.13	0.33	10.09	25,464

Notes: Mean voting-power share, approval indicator, and author-choice indicator by choice rank. Panel A uses all proposals in the active choice-level data after release guards; Panel B restricts the same statistics to proposals with up to three choices. Author-choice means omit rows without reconstructed author-choice information.

5 Disentangling Biases: A Regression Approach

The descriptive results show that the three candidate biases are tightly entangled. Table 3 makes this overlap explicit: first-position choices receive most voting power, but first-position choices are also much more likely to be approving choices and author-selected choices. Raw averages therefore cannot tell whether an author-selected choice attracts voting power because it is first, because it approves the proposal, because it was selected by the author, or because these indicators co-occur. We use regression models to separate these conditional associations while preserving the observational interpretation of the analysis.

5.1 Model Design and Samples

Model The unit of analysis are proposal's choices and the outcome is the share of voting-power of each choice (bounded between zero and one). We estimate fractional-logit models with a quasi-binomial logit link, which models fractional outcomes without converting the voting-power share into a binary outcome. The reported coefficient estimates are on the logit scale; substantive interpretation therefore relies on the response-scale average marginal effects (AMEs) reported in the robustness tables.

Table 4: Robustness specification key

#	Short name	Dimension	Definition	N
1	Main	Baseline	Single-preference proposals: single-choice and basic voting systems.	293,816
2	Vote-count	Outcome	Vote-count share (number of votes divided by total proposal votes) as outcome.	293,816
3	Auth-voted	Participation	Restricted to proposals where the author voted.	172,171
4	Low-part.	Participation	Excludes proposals with fewer than two votes.	221,774
5	Vote-control	Participation	Adds log total proposal votes.	293,816
6	TVL-25	TVL	Restricted to Top 25 Ethereum-TVL DAOs matched against DefiLlama ranking.	11,410
7	TVL-50	TVL	Restricted to Top 50 Ethereum-TVL DAOs matched against DefiLlama ranking.	14,393
8	TVL-100	TVL	Restricted to Top 100 Ethereum-TVL DAOs matched against DefiLlama ranking.	15,468
9	Proposal FE	Fixed effects	Uses proposal fixed effects.	293,816
10	No top-5	Space influence	Excludes the five largest contributing spaces.	272,743
11	Expanded	Proposal type	Includes approval and ranked-choice voting systems.	308,582
12	Full	Proposal type	Includes all voting systems, including weighted and quadratic.	398,216

Notes: Variations 2-10 are variations of the main single-preferences sample; variations 11-12 expands the sample to multi-preference proposals. The TVL samples contain the DAOs matched against the Top 100 Decentralized Finance (DeFi) protocols ranked by DefiLlama (TVL refers to the median daily Ethereum total value locked by DeFi protocols over the study period); accordingly TVL-100 contains less than 100 DAOs because not every protocol can be matched (see Appendix D.4 for details).

Specification We control for the three bias indicators (author selection, approving stance, choice position) plus a set of covariates: relative proposal date, number of choices, author-choice visibility among the first six interface positions,[6] relative choice-text length, and voting-system in use. In addition, because the author of a proposal may or may not have voted, we include an indicator that it is equal to one if the proposal author voted, zero otherwise. The main fixed effect is the DAO space, which absorbs persistent differences across governance communities (robustness checks test also proposal fixed effects to compare choices within the same proposal). Standard errors are clustered by space throughout because choices and proposals from the same DAO are not independent observations.

Main sample The primary regression sample uses choices from all single-preference proposals, where the mapping from a ballot to a choice is most direct.

[6] Six is a constraint of the user interface of Snapshot.

Table 5: Nested fractional-logit specifications for the main single-preference sample

	(1)	(2)	(3)	(4)	(5)	(6)
Author choice	3.347***	3.347***	3.348***	3.348***	3.352***	2.073***
	(0.086)	(0.086)	(0.087)	(0.087)	(0.086)	(0.088)
Approving choice	1.750***	1.750***	1.754***	1.754***	1.751***	1.768***
	(0.141)	(0.141)	(0.140)	(0.140)	(0.139)	(0.139)
Choice rank	-0.529***	-0.528***	-0.517***	-0.517***	-0.511***	-0.510***
	(0.038)	(0.038)	(0.039)	(0.039)	(0.039)	(0.039)
Choice rank squared	0.003***	0.003***	0.003***	0.003***	0.003***	0.003***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Author voted	-1.546***	-1.546***	-1.546***	-1.524***	-1.526***	-0.949***
	(0.039)	(0.039)	(0.039)	(0.037)	(0.037)	(0.039)
Relative proposal date		-0.000	-0.000	-0.000	-0.000	-0.000
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Number of choices			-0.029**	-0.029**	-0.030**	-0.030**
			(0.009)	(0.009)	(0.009)	(0.009)
Author choice in first six				-0.026+	-0.026+	-0.667***
				(0.013)	(0.013)	(0.045)
Author choice x first-six visibility						1.412***
						(0.091)
Relative choice text length					-0.118*	-0.122*
					(0.059)	(0.059)
Voting-type controls	Yes	Yes	Yes	Yes	Yes	Yes
N	293816	293816	293816	293816	293816	293816

Notes: Coefficient estimates are on the logit scale with DAO-space clustered standard errors in parentheses. Significance markers: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Robustness experiments We use alternative sample specifications and model variations to verify the robustness of our findings. Table 4 defines our experiments, the short names used to identify them in the robustness tables, and reports the fitted row count for each of them.

5.2 Results

Table 5 reports nested fractional-logit specifications for the main single-preference sample. Across all six specifications, the author-choice coefficient remains stable and is the strongest predictor of voting-power share among the three biases, followed by approving stance and choice rank. Because these coefficients are estimated on the logit scale, we turn to average marginal effects (AMEs) for their substantive interpretation: AMEs express how many percentage points a choice's voting-power share changes, on average, when the corresponding indicator varies, holding other covariates fixed. Figure 5 reports these AMEs for our Main specification alongside all robustness checks. In the Main specification, an author-selected choice gains 58.8 percentage points of voting-power share relative to an otherwise similar choice, compared with 27.1 points for having an approving stance, and 7.7 points for being listed first rather than second. The remaining robustness specifications recover the same ordering, with the author choice being the largest, followed by approving stance

Table 6: Author-specific diagnostics for author-choice AMEs

			Author-Position Split		Author-Stance Split	
Author's VP	N	Pooled	First pos.	Non-first	Approve	Non-Approve
Included	122,582	47.92	59.47	36.58	56.16	46.38
Removed	122,582	32.56	45.06	20.72	46.26	30.66

Notes: the sample includes all proposals in which the author voted and with at least one more vote casted with positive voting power. Cells report response-scale author-choice AMEs in percentage points.

and position, confirming that this ranking does not depend on any single sample restriction or model choice.

To better understand the role of the author bias, we conducted an additional diagnostics test in which we subtract the author's own voting power from the final voting power share, also controlling for author-choice positioning and approving-choice status. For this purpose, we derive a new sample from *Main*, keeping only proposals in which the author voted and with at least one additional vote with positive voting power was cast. We report three specifications on this sample: (i) Pooled, which includes all selected proposal choices; (ii) Author $\times$ Position, which splits the sample into 'author choice in first position' vs. 'author choice elsewhere'; and (iii) Author $\times$ Stance, which splits it into 'author choice with Approve stance' vs. 'author choice with any other stance'.

Table 6 reports the results. In the top row, we report the AME values when the author's voting power is included; in the bottom row, we report the AME values when the author's voting power is excluded. On the pooled sample, the observed author-choice AME is 47.92 percentage points and falls to 32.56 percentage points after subtracting the author's own support, which suggests that the author-choice pattern is not mechanically reducible to authors voting for their own choices. The position split shows that the association is present for both first-position and non-first author choices, albeit weaker. The approve split likewise shows positive author-choice AMEs for both approving and non-approving author choices, with the latter showing a weaker association as expected.

Taken together, the regression evidence refines the descriptive pattern from Section 4. Across all specifications, author-selected choices show a large and robust conditional association with voting-power shares, with the association remaining visible even after separating author participation from author choice, and after removing author voting power. This evidence, however, remains a descriptive author-choice association, and not a causal claim about persuasion or manipulation.

6 Conclusion

This paper examines how voting power is distributed across choices in the governance proposals of decentralized autonomous organizations (DAOs) on the Snapshot voting platform and the extent to which the observed high approval rates can be explained by three biases: *position*, capturing the

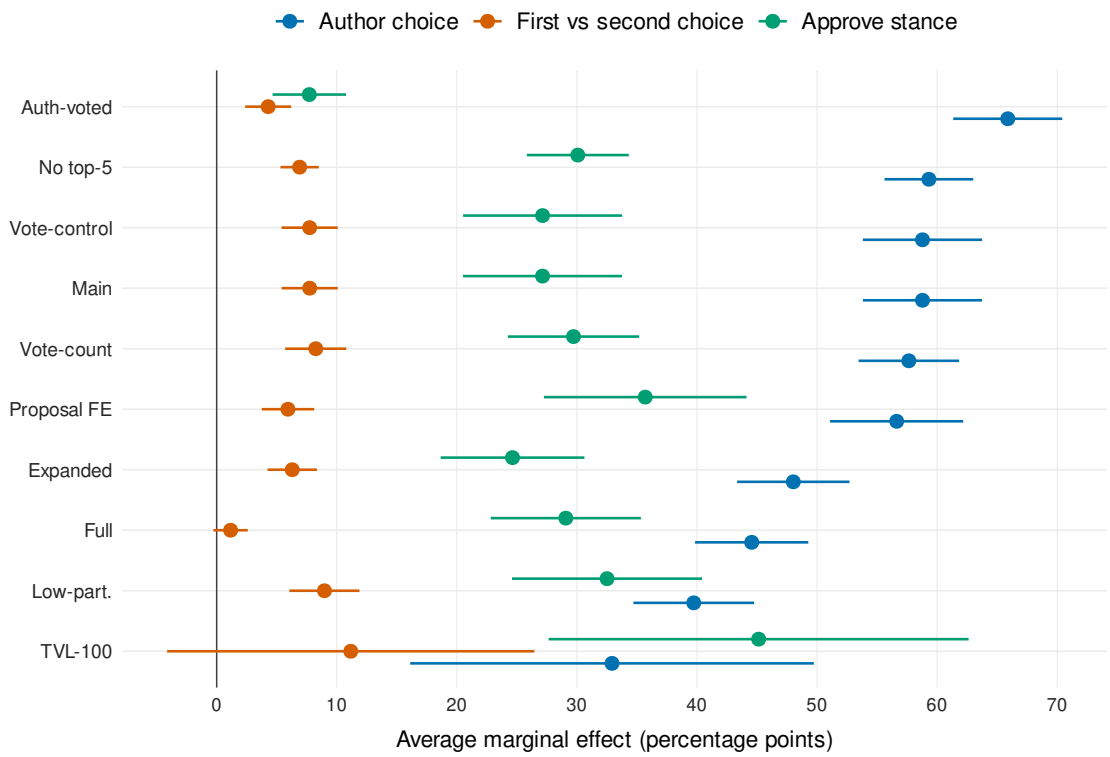


Fig. 5: **Three-bias average marginal effects (AMEs) across main sample and robustness specifications**. Points show response-scale AMEs in percentage points; horizontal intervals are Bonferroni FWER-adjusted across the 36 shared AMEs. *Main* corresponds to the primary specifications; others are robustness checks. Across all ten robustness specifications, the author-choice AME exceeds both the position effect (first vs. second choice) and the approval-stance effect. While percentages vary across specifications, signs and relative size remain consistent, confirming that author selection is not an artifact of any single sample restriction or model variant. *TVL-25* and *TVL-50* rows omitted for readability; estimates are comparable to *TVL-100*, see Table 17 in the Appendix for details.

importance of the list ordering, *approval*, capturing the systematic impact of a choice's stance, and *author*, capturing reputation or social influence signals. Across the observed data, three regularities stand out. First, choices listed in the first position receive substantially more voting power than later choices, with an advantage of 7.7 percentage points; second, approving choices are also favored, with an estimated 27.1 percentage points advantage; and third, author-selected choices show the largest conditional association with voting-power shares, with a a 58.8 percentage-point average marginal effect in our main specification.

The robustness analyses sharpen the interpretation of the author-choice result. The association remains large across multiple samples and specifications. Notably, the author bias also remains visible when removing from the outcome the author's own voting power: the author-choice AME falls from 47.92 to 32.56 percentage points, but does not disappear. Notably, restricting the sample to the top DeFi protocols reduces considerably the effect of the author bias so that it becomes smaller than the approve bias, even if the difference is not significant; however, this sample is notably smaller and the estimates might be noisier. This result is consistent with the reduction of author bias once low participation proposals are removed.

The author bias pattern is consistent with more than one mechanism. Author-selected choices may receive support because authors cast voting power for their own preferred options, because their choices convey information to later voters, because they are strategically placed in prominent positions, or because authors select choices that are already more likely to attract support. The present design distinguishes these associations more cleanly than raw averages, but it does not isolate a single causal channel.

Implications. The findings point to a practical transparency issue in DAO governance. If author-selected choices systematically receive more voting power, voters and observers may need clearer information about who created a proposal, which option the author selected, and how much of the final score comes from the author's own voting power. This does not imply that author influence is improper. Proposal authors often have relevant expertise and may legitimately signal which option implements the proposal as intended. Still, the concentration of voting power around author-selected choices matters for accountability, especially when authors are pseudonymous, affiliated with organizations, or connected to multiple governance spaces. The result is also relevant to concerns that large governance token holders and other vested contributors can shape outcomes in ways that ordinary voters may not easily observe [17]. These findings also make interface choices analytically important. Snapshot ballots present choices in an ordered list, expose voting activity under many configurations, and allow observers to infer highly uneven support before treating an outcome as broad consensus. For governance design, that means choice order, author-choice visibility, and author-vote visibility should be documented and, where possible, made auditable rather than treated as neutral interface details. For empirical interpretation, highly skewed voting outcomes should not automatically be read as unambiguous preference aggregation: they may also

reflect proposal screening, prominent author signals, list order, and the public accumulation of voting power during the vote.

Limitations. The analysis is observational and therefore quantifies systematic associations rather than causal effects. Consistent with the positive-outcome skew discussed in Section 2, high approval rates may partly reflect proposal screening before formal voting: proposals that fail informal "temperature checks" or forum discussions may be abandoned or revised before reaching Snapshot. Our dataset therefore observes only proposals that survive this filtering, so lopsided approval rates may partly reflect a survival process rather than how voters treat proposals once balloted. This boundary cannot be removed by simply collecting more pre-vote material, because the outcome variable, voting-power share, as well as choices, choice order, and author selection are defined only for formal Snapshot ballots. The residual author-choice association after removing author voting power may reflect information, coordination, ballot design, social influence, or unmeasured proposal quality. Social-influence interpretations also depend on vote and tally visibility during the voting period; privacy-preserving mechanisms such as the Shutter plugin may follow different dynamics, but the current Shutter evidence is not yet strong enough to identify that channel. Future work should combine Snapshot data with forum discussions, proposal histories, author vote timing, and on-chain records to separate these mechanisms more directly, while treating proposals that never reach a vote as outside the ballot-level regression framework used here.

Overall, the paper provides evidence that DAO voting outcomes are not only shaped by formal token-weighted aggregation. They are also associated with interface position, stance structure, and especially the choices selected by proposal authors. Those patterns do not by themselves establish misconduct or manipulation, but they show that descriptive features of the ballot and of the proposal author are central to understanding how voting power accumulates in decentralized governance.

References

1. Andrés Abeliuk, Gerardo Berbeglia, Pascal Van Hentenryck, Tad Hogg, and Kristina Lerman. Taming the unpredictability of cultural markets with social influence. In *Proceedings of the 26th international conference on world wide web*, pages 745–754, 2017.
2. Hayden Adams, Noah Zinsmeister, Moody Salem, River Keefer, and Dan Robinson. Uniswap v3 core. Technical report, Uniswap, 2021. Available at `https://uniswap.org/whitepaper-v3.pdf`.
3. Aman Agarwal, Xuanhui Wang, Cheng Li, Michael Bendersky, and Marc Najork. Addressing trust bias for unbiased learning-to-rank. In *The World Wide Web Conference*, pages 4–14, 2019.
4. Raphael Auer, Bernhard Haslhofer, Stefan Kitzler, Pietro Saggese, and Friedhelm Victor. The technology of decentralized finance (DeFi). *Digital Finance*, August 2023. `doi:10.1007/s42521-023-00088-8`.
5. Stefano Balietti. nodeGame: Real-time, synchronous, online experiments in the browser. *Behavior Research Methods*, 49(5):1696–1715, 2017.
6. Stefano Balietti, Can Celebi, and David Tercero-Lucas. From crypto to nfts: Identifying the new wave of digital investors. *International Review of Financial Analysis*, 104:104172, 2025.
7. Stefano Balietti, Pietro Saggese, Stefan Kitzler, and Bernhard Haslhofer. Slaying the dragon: The quest for democracy in decentralized autonomous organizations (daos). *arXiv preprint arXiv:2511.09263*, 2025.
8. Abhijit V. Banerjee. A simple model of herd behavior. *The Quarterly Journal of Economics*, 107(3):797–817, 1992. `doi:10.2307/2118364`.
9. Tom Barbereau, Reilly Smethurst, Orestis Papageorgiou, Johannes Sedlmeir, and Gilbert Fridgen. Decentralised finance's timocratic governance: The distribution and exercise of tokenised voting rights. *Technology in Society*, page 102251, April 2023. URL: `https://www.sciencedirect.com/science/article/pii/S0160791X23000568`, `doi:10.1016/j.techsoc.2023.102251`.
10. Tom Josua Barbereau, Reilly Smethurst, Orestis Papageorgiou, Alexander Rieger, and Gilbert Fridgen. Defi, not so decentralized: The measured distribution of voting rights. In *Proceedings of the Hawaii International Conference on System Sciences 2022*, page 10, 2022.
11. Sushil Bikhchandani, David Hirshleifer, and Ivo Welch. A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy*, 100(5):992–1026, 1992. `doi:10.1086/261849`.
12. Ethan Bito, Yongli Ren, and Estrid He. Evaluating position bias in large language model recommendations. *arXiv preprint arXiv:2508.02020*, 2025.
13. Eric F. Buddensiek and Paul P. Momtaz. Information cascades in decentralized autonomous organizations (daos), 2025. SSRN working paper. `doi:10.2139/ssrn.5258156`.
14. Keith Burghardt, Emanuel F Alsina, Michelle Girvan, William Rand, and Kristina Lerman. The myopia of crowds: Cognitive load and collective evaluation of answers on stack exchange. *PloS one*, 12(3):e0173610, 2017.
15. Keith Burghardt, Tad Hogg, Raissa D'Souza, Kristina Lerman, and Marton Posfai. Origins of algorithmic instabilities in crowdsourced ranking. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW2):1–20, 2020.
16. Keith Burghardt, Tad Hogg, and Kristina Lerman. Quantifying the impact of cognitive biases in question-answering systems. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 12, 2018.

17. Vitalik Buterin. Moving beyond coin voting governance, 2021. Available at https://vitalik.ca/general/2021/08/16/voting3.html.
18. L Elisa Celis, Peter Krafft, and Nathan Kobe. Sequential voting promotes collective discovery in social recommendation systems. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 10, pages 42–51, 2016.
19. Andrew Collins, Dominika Tkaczyk, Akiko Aizawa, and Joeran Beel. A study of position bias in digital library recommender systems. *arXiv preprint arXiv:1802.06565*, 2018.
20. Compound. Improvement proposal: Compound v2 to v3 migration phase 1, 2023. Available at: https://compound.finance/governance/proposals/152.
21. Lin William Cong, Daniel Rabetti, Charles CY Wang, and Yu Yan. Centralized governance in decentralized organizations. *Available at SSRN 5168660*, 2025.
22. ConstitutionDAO. Documentation. Technical report, ConstitutionDAO, 2021. Available at https://www.constitutiondao.com/.
23. Nick Craswell, Onno Zoeter, Michael Taylor, and Bill Ramsey. An experimental comparison of click position-bias models. In *Proceedings of the 2008 international conference on web search and data mining*, pages 87–94, 2008.
24. Curve. Improvement proposal: reduction to crv token inflation, 2020. Available at: https://gov.curve.fi/t/discussion-reduction-to-crv-inflation/851.
25. Himel Dev, Karrie Karahalios, and Hari Sundaram. Quantifying voter biases in online platforms: An instrumental variable approach. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW):1–27, 2019.
26. Maya Dotan, Aviv Yaish, Hsin-Chu Yin, Eytan Tsytkin, and Aviv Zohar. The Vulnerable Nature of Decentralized Governance in DeFi. Technical report, The Hebrew University, August 2023. arXiv:2308.04267 [cs] type: article. URL: http://arxiv.org/abs/2308.04267, doi:10.48550/arXiv.2308.04267.
27. Youssef Faqir-Rhazoui, Javier Arroyo, and Samer Hassan. A comparative analysis of the platforms for decentralized autonomous organizations in the ethereum blockchain. *Journal of Internet Services and Applications*, 12(1):9, 2021. doi:10.1186/s13174-021-00139-6.
28. Rainer Feichtinger, Robin Fritsch, Yann Vonlanthen, and Roger Wattenhofer. The hidden shortcomings of (d) aos–an empirical study of on-chain governance. *arXiv preprint arXiv:2302.12125*, 2023.
29. Robin Fritsch, Marino Müller, and Roger Wattenhofer. Analyzing voting power in decentralized governance: Who controls daos?, 2022. arXiv:2204.01176.
30. Mitchell Goldberg and Fabian Schär. Metaverse Governance: An Empirical Analysis of Voting Within Decentralized Autonomous Organizations. Technical Report 4310845, Rochester, NY, December 2022. URL: https://papers.ssrn.com/abstract=4310845, doi:10.2139/ssrn.4310845.
31. Pamela Grimm. Social desirability bias. *Wiley international encyclopedia of marketing*, 2010.
32. Samer Hassan and Primavera De Filippi. Decentralized autonomous organization. *Internet Policy Review*, 10(2), 2021. doi:10.14763/2021.2.1556.
33. Tad Hogg and Kristina Lerman. Disentangling the effects of social signals. *arXiv preprint arXiv:1410.6744*, 2014.
34. Johannes Rude Jensen, Victor von Wachter, and Omri Ross. How decentralized is the governance of blockchain-based finance: Empirical evidence from four governance token distributions, 2021. arXiv:2102.10096.

35. Thorsten Joachims, Laura Granka, Bing Pan, Helene Hembrooke, and Geri Gay. Accurately interpreting clickthrough data as implicit feedback. In *28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2005*, pages 154–161, 2005.
36. Stefan Kitzler, Stefano Balietti, Pietro Saggese, Bernhard Haslhofer, Markus Strohmaier, et al. The governance of decentralized autonomous organizations: A study of contributors' influence, networks, and shifts in voting power. In *Financial Cryptography and Data Security 2024*. IFCA, 2024.
37. Jonathan G. S. Koppell and Jennifer A. Steen. The effects of ballot position on election outcomes. *The Journal of Politics*, 66(1):267–281, 2004. doi:10.1046/j.1468-2508.2004.00151.x.
38. Jon A. Krosnick and Duane F. Alwin. An evaluation of a cognitive theory of response-order effects in survey measurement. *Public Opinion Quarterly*, 51(2):201–219, 1987.
39. Robert Leshner and Geoffrey Hayes. Compound: The money market protocol. Technical report, Compound, 2019. Available at https://bit.ly/3ioWOjW.
40. Yuyang Liang. Knowledge sharing in online discussion threads: What predicts the ratings? In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, pages 146–154, 2017.
41. Chang Liu, Yixin Wang, and Moontae Lee. Counterfactual voting adjustment for quality assessment and fairer voting in online platforms with helpfulness evaluation. *arXiv preprint arXiv:2506.21362*, 2025.
42. J. Lorenz, H. Rauhut, F. Schweitzer, and D. Helbing. How social influence can undermine the wisdom of crowd effect. *Proceedings of the National Academy of Sciences (PNAS)*, 108(22):9020–9025, 2011.
43. MakerDAO. The maker protocol: Makerdao's multi-collateral dai (mcd) system. Technical report, MakerDAO, 2020. Available at https://makerdao.com/en/whitepaper.
44. Emaad Manzoor, George H Chen, Dokyun Lee, and Michael D Smith. Influence via ethos: On the persuasive power of reputation in deliberation online. *Management Science*, 70(3):1613–1634, 2024.
45. Francesco Marolla, Angelica Maineri, Jacopo Tagliabue, and Giovanni Cassani. Voting, fast and slow: ballot order and likeability effects in the five-star movement 2012 online primary election. *Contemporary Italian Politics*, 16(3):266–283, 2024.
46. Marc Meredith and Yuval Salant. On the causes and consequences of ballot order effects. *Political Behavior*, 35(1):175–197, 2013. doi:10.1007/s11109-011-9189-2.
47. Johnnatan Messias, Vabuk Pahari, Balakrishnan Chandrasekaran, Krishna P. Gummadi, and Patrick Loiseau. Understanding blockchain governance: Analyzing decentralized voting to amend defi smart contracts, 2024. arXiv:2305.17655.
48. Lev Muchnik, Sinan Aral, and Sean J Taylor. Social influence bias: A randomized experiment. *Science*, 341(6146):647–651, 2013.
49. Matthias Nadler and Fabian Schär. Decentralized Finance, Centralized Ownership? An Iterative Mapping Process to Measure Protocol Token Distribution. *arXiv:2012.09306 [cs, econ, q-fin]*, December 2020. arXiv: 2012.09306. URL: http://arxiv.org/abs/2012.09306.
50. Aditya Pal and Scott Counts. Identifying topical authorities in microblogs. In *Proceedings of the fourth ACM international conference on Web search and data mining*, pages 45–54, 2011.
51. Sharoda A Paul, Lichan Hong, and Ed H Chi. Who is authoritative? understanding reputation mechanisms in quora. *arXiv preprint arXiv:1204.3724*, 2012.
52. Paul Resnick, Ko Kuwabara, Richard Zeckhauser, and Eric Friedman. Reputation systems. *Communications of the ACM*, 43(12):45–48, 2000. doi:10.1145/355112.355122.

53. Romain Rossello. Blockholders and strategic voting in daos' governance. *Available at SSRN 4706759*, 2024.
54. Romain Rossello, Stefano Balietti, Stefan Kitzler, and Pietro Saggese. Voters' preferences and the willingness to pay for voting rights: New evidence from decentralized finance. *Available at SSRN 6313439*, 2026.
55. M.J. Salganik, P.S. Dodds, and D.J. Watts. Experimental study of inequality and unpredictability in an artificial cultural market. *Science*, 311(5762):854–856, 2006.
56. Snapshot. Documentation. Technical report, Snapshot, 2023. Available at `https://docs.snapshot.org/`.
57. Greg Stoddard. Popularity dynamics and intrinsic quality in reddit and hacker news. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 9, pages 416–425, 2015.
58. Xiaotong Sun, Charalampos Stasinakis, and Georigios Sermpinis. Decentralization illusion in decentralized finance: Evidence from tokenized voting in makerdao polls, 2023. `arXiv:2203.16612`.
59. Sushiswap. Documentation. Tech. rep., Sushiswap, 2022. Available at `https://docs.sushi.com/`.
60. Yla R Tausczik and James W Pennebaker. Predicting the perceived quality of online mathematics contributions from users' reputations. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 1885–1888, 2011.
61. Uniswap. Improvement proposal: Making protocol fees operational, 202. Available at: `https://gov.uniswap.org/t/making-protocol-fees-operational/21198`.
62. Uniswap. Improvement proposal: deploy uniswap v3 on bnb chain, 2022. Available at: `https://gov.uniswap.org/t/rfc-update-deploy-uniswap-v3-1-0-3-0-05-0-01-on-bnb-chain-binance/19734`.
63. Uniswap. Improvement proposal: Uniswap dao can help 80k uni to turkiye to relief after the big earthquake disaster, 2023. Available at: `https://gov.uniswap.org/t/governance-proposal-uniswap-dao-can-help-80k-uni-to-turkiye-to-relief-after-the-big-earthquake-disaster/20758`.
64. Peiyi Wang, Lei Li, Liang Chen, Zefan Cai, Dawei Zhu, Binghuai Lin, Yunbo Cao, Lingpeng Kong, Qi Liu, Tianyu Liu, et al. Large language models are not fair evaluators. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 9440–9450, 2024.
65. Qin Wang, Guangsheng Yu, Yilin Sai, Caijun Sun, Lam Duc Nguyen, Sherry Xu, and Shiping Chen. An empirical study on snapshot daos. *ArXiv*, abs/2211.15993, 2022. URL: `https://api.semanticscholar.org/CorpusID:254069759`.
66. Shuai Wang, Wenwen Ding, Juanjuan Li, Yong Yuan, Liwei Ouyang, and Fei-Yue Wang. Decentralized autonomous organizations: Concept, model, and applications. *IEEE Transactions on Computational Social Systems*, 6(5):870–878, 2019.
67. Lukas Weidener and Bence Lukács. The decentralization of science and the emergence of decentralized science. *Available at SSRN 5094597*, 2025.
68. Aviv Yaish, Svetlana Abramova, and Rainer Böhme. Strategic vote timing in online elections with public tallies. *arXiv preprint arXiv:2402.09776*, 2024.
69. Zimo Yang, Zi-Ke Zhang, and Tao Zhou. Anchoring bias in online voting. *Europhysics Letters*, 100(6):68002, 2012.

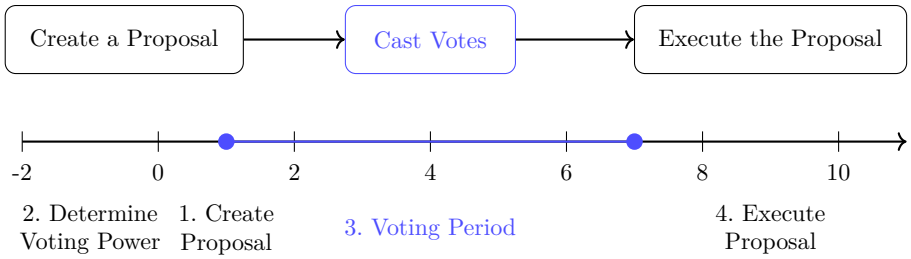


Fig. 6: **Timeline of a DAO governance process**. The decision-making process can be divided into three time windows: the pre-voting period, the voting period, and the post-voting period.

A Decentralized Autonomous Organizations

Figure 6 summarizes the governance sequence used throughout the empirical analysis. A proposal is created, voting power is determined by the rules of the relevant DAO space, token holders cast votes during the voting window, and the winning choice can then be executed or used as a governance signal.

B Data Processing

The original dataset, obtained from Snapshot.org, contains 57 626 386 individual votes spread across 223 760 governance proposals from 32 995 DAO spaces, for a total of 653 913 choices, over the period Oct 2020 to Nov 2023. Below, we discuss the various data cleaning procedures that we applied.

B.1 Initial Processing

We conducted a series of sanity checks and cleaning procedures in two stages. At the proposal level, following [36], we removed proposals labeled as 'pending' or 'invalid' and retained those marked as 'final'. We also eliminated proposals with nonsensical titles, template text in the body, or template text in the title, such as "Lorem ipsum dolor sit amet, consectetur adipiscing elit [...]", "new title", "new proposal", or "Title of your proposal goes here". We further excluded proposals with less than two choices and proposals marked by Snapshot's "flagged" field.

At the vote level, we identified invalid votes, i.e., votes not conforming to the choices admitted by the proposal. These include plain malformed votes, such as votes containing an address rather than an integer, as well as votes containing an integer pointing to a non-existing choice, such as index 4 when only three choices are available. Snapshot.org marks these votes as invalid,[7] while the API still exposes them. We excluded malformed votes and votes pointing to unavailable choices,

[7] See for instance: `https://snapshot.org/#/testsnap.eth/proposal/0x1b1c6e19edb5358110aa20e6e261d92b22bc2422925d759b5968e190f33ecbf5`.

and recorded duplicate-ballot and invalid-vote indicators as diagnostics. The cleaning code records all cleaning decisions for reproduction.

After filtering to final proposals, the vote-merge validation starts from 169 689 proposals. The resulting analytic dataset contains 19 201 spaces, with 127 281 proposals for a total of 398 218 choices and 50 759 796 favorite-mode ballot mass. This amounts to an average of 3.13 choices per proposal ($\pm$ 8.29), with a median of 2; spaces on average have 6.63 proposals ($\pm$ 38.71), with a median of 1. The final dataset was checked to be free from invalid or duplicated votes and from all proposals that were flagged with issues by the Snapshot team.

We classify spaces as 'mature' when they fulfill all of the following three criteria from [36]: (i) at least five followers, (ii) more than two voters in any proposal, and (iii) more than one proposal. All spaces in the resulting analytic dataset satisfy these maturity criteria used in the downstream analysis. We additionally excluded 83 proposals marked by Snapshot's "flagged" field, corresponding to 58.04% of the 143 raw proposals with this flag. We evaluate the quality of these procedures by looking at the retention of mature spaces. Encouragingly, we find that only 5% of the eliminated spaces were classified as mature.

B.2 Processing Votes

The primary voting-power outcome in the paper is choice-level voting-power share. Most proposals (93%) follow a single-preference choice model (see Table 2 in main text), which makes the computation of the voting power share straightforward (i.e., choice's voting power divided the total voting power in a proposal).

However, in some voting systems (as also discussed in Section 2), users can express preferences for multiple choices with a single vote. For example, in weighted choice models, users can select multiple options and assign weights reflecting their relative preference. In these cases, we need to uniformly map users' raw votes to a set of expressed choices for vote-count diagnostics. To do so, we split multiple-preference ballots into one or multiple choice-level rows while conserving each ballot's unit mass.

We test two diagnostic mappings for assigning ballot mass to choices in the presence of multiple preferences: (i) favorite-choice, and (ii) support-fraction. The former assigns each ballot's unit mass to the favorite choice or choices; approval ballots split mass equally across the distinct approved choices, and weighted or quadratic ballots keep the maximum positive-weight choices and split ties equally. The latter divides ballot mass across supported choices; for weighted and quadratic ballots, positive weights are normalized across all selected choices. Table 7 reports how users' raw votes are mapped to a set of expressed choices in different multiple-preference voting systems.

Both mappings may lead to more choice-level rows than ballot rows in multiple-choice voting systems, but neither duplicates a ballot's mass as a full vote for every selected choice. The vote-expansion procedure processed 148 056 proposals and produced 267 437 favorite-choice rows, with 13 554 favorite-choice parse-error rows retained as diagnostics.

The support-fraction mapping also defines the no-author-voting-power sensitivity used in the regression analyses. When information on author votes is available, the adjusted score removes the portion of voting power attributable to the proposal author from each choice supported by the author. Specifically, the author's total voting power is allocated across all author-supported choices in proportion to the author's support for those choices, rather than being subtracted entirely from a single favorite choice. The proposal-level choice shares are then recomputed using the adjusted voting power.

Practically, between the two mapping systems the differences are small and they occur only for weighted and quadratic proposals, single-choice, basic, ranked-choice, and approval have zero mapping differences. Quantitatively, we register differences across the mappings in 2,721 weighted proposals and 917 quadratic proposals for a total of 30,929 proposal-choice pairs. Only 8,263.57 out of 54,978,057 valid ballot mass, about 0.015% overall, changes across the two systems, with a median redistributed proposal mass of 3.47 percentage points, 95th percentile is 40.0 percentage points, max is 87.13 percentage points.

Across all the variations in Table 4, author choice is defined with the favorite-choice mapping. For variations 1-11, this either is the only relevant mapping or is practically identical to support-fraction under the included voting systems. In variation 12 (Full), this is a substantive choice: weighted and quadratic ballots are included, but the author-choice regressor remains favorite-choice based.

In Table 6, the author-choice indicator is still defined using the favorite-choice mapping: a choice is treated as author-selected if it belongs to the author's favorite-choice set. The support-fraction mapping is used only to construct the no-author-voting-power outcome. For this sensitivity, we subtract the author's own voting power from each choice in proportion to the author's support fraction—i.e., the author voting power times the author's support fraction—and then recompute proposal-level choice shares. Thus, the table keeps the author-choice contrast fixed while comparing observed voting-power shares to shares after removing the author's own voting power.

C Choice Stance Detection

The choices associated with a proposal can be expressed with simple binary responses, numbers, strings of text, or other arbitrary content, and may include indications to approve or reject its content, or more complex options. We thus try to classify each choice of all proposals to a simplified *stance*, defined as the intended meaning or position expressed by a voting choice within a proposal.

To classify the stance of a choice, we need to first make an important distinction between proposals having a *direct effect* on the governance of its organization

We assign a stance of "Reject" or "Approve" only to choices of proposals that are directly rejecting or approving governance changes; often cases in such proposals there exists a third choice to abstain from making a decision, which we label "Abstain" (such a choice helps reaching the quorum, without expressing a preference).

Table 7: Multi-preference vote processing

Voting system	**Raw ballot**	**Favorite-choice rows choice: mass**	**Support-fraction rows choice: mass**
Basic	1 *rows: 1, mass: 1*	1: 1 *rows: 1, mass: 1*	1: 1 *rows: 1, mass: 1*
Single-choice	2 *rows: 1, mass: 1*	2: 1 *rows: 1, mass: 1*	2: 1 *rows: 1, mass: 1*
Weighted	{1: 50, 2: 50} *rows: 1, mass: 1*	1: 1/2 2: 1/2 *rows: 2, mass: 1*	1: 1/2 2: 1/2 *rows: 2, mass: 1*
Quadratic	{1: 3, 2: 1} *rows: 1, mass: 1*	1: 1 *rows: 1, mass: 1*	1: 3/4 2: 1/4 *rows: 2, mass: 1*
Ranked-choice	[4, 2, 3, 1] *rows: 1, mass: 1*	4: 1 *rows: 1, mass: 1*	4: 1 *rows: 1, mass: 1*
Approval	[1, 2, 4] *rows: 1, mass: 1*	1: 1/3 2: 1/3 4: 1/3 *rows: 3, mass: 1*	1: 1/3 2: 1/3 4: 1/3 *rows: 3, mass: 1*

Notes: initially, a raw ballot is a single row in the dataset carrying unit mass (*rows: 1, mass: 1*); the mapping expands it into one or more rows that always sum to the same unit mass, so no ballot is double-counted even when the number of rows grows. The favorite-choice mapping assigns mass to the maximum-weight choice(s) (ties split equally) and equally splits approval ballots across all approved choices. The support-fraction mapping normalizes positive weights across all selected choices for weighted and quadratic ballots, splits approval ballots equally, and keeps the top-ranked choice only for ranked-choice ballots.

Finally, we assign the label "Other" to all the choices of proposals that do not have a direct governance effect, and more broadly to all the choices that cannot be directly classified as either "Reject", "Approve", or "Abstain."

Importantly, a proposal must offer at least one approval and one refusal choice, so that any choice could be labeled as either "Approve" or "Reject"; otherwise, all choices are labeled as "Other." For instance, simply selecting an alternative between a number of choices does not follow an approve/reject pattern, because there is no possibility to object to the proposal altogether; in this case, all choices will be labeled as "Other." However, if there is one choice rejecting all the alternatives, we mark this choice as "Reject", and the other as "Approve" (more "Reject" and "Abstain" choices in the same proposals are also allowed). Before this recoding, we registered 4,274 reject-only and 1,945 approve-only proposals that we normalized to "Other" (see Table S 11).

Having stated our classification goal, we implemented it using rule-based heuristics and validated via large language models and human evaluation, as described below.

C.1 Heuristics

The heuristic procedure consists of two substantive classification steps followed by a normalization and exception-handling step. First, a registry-backed rule classifier assigns an initial stance to each choice using exact, prefix, suffix, high-confidence-contains, and low-confidence-contains patterns. Second, we apply targeted contextual checks for cases that cannot be reliably resolved from isolated choice text, including negated action labels, stance-bearing residual options, and manually inspected emoji choices. Finally, we reconcile known exceptions and enforce the final proposal-level consistency criteria, including space- and proposal-level overrides, per-proposal stance-sequence overrides, and the normalization of undocumented one-sided Approve/Reject proposals to *Other*.

Registry-backed rule classifier. The general keyword rules are defined in the stance-pattern registry used by the classifier. The registry starts from conservative stance-specific exact terms and expands them into prefix, suffix, and contains tiers, with short or ambiguous tokens kept exact-only. Tables 8–10 summarize the registry tiers and their resolution rules by stance.

Table 8: Reject stance heuristic registry

Tier	Rule	Terms
Exact	Trimmed lower-case choice equals the canonical term.	English: againist, against, aganist, decline, declined, denied, deny, disagree, disagreed, disapporve, disapprove, disapproved, dissaprove, dissent, do not, do not accept, do not approve, do not proceed

Table 8: Reject stance heuristic registry (continued)

Tier	Rule	Terms
		English: do not support, do nothing, don't, don't accept, don't agree, don't approve, don't support, dont agree, dont support, dont't proceed, i deny, i disagree, i dissent, i do not accept, i do not agree, i do not approve, i do not support, i don't accept English: i don't agree, i don't approve, i don't support, i dont agree, i dont support, i refuse, i reject, ino, n, nae, nah, nay, negative, negative vote, no, no support, nope, not accept English: not agree, not approve, not approved, not in favor, not support, oppose, opposed, refuse, refused, reject, rejected Non-English: contra, no apruebo, non, non approvo, não, нет, 不会, 不可以, 不同意, 不支持, 不是, 不能, 反对, 反对, 否, 没有
Prefix	Left-trimmed choice starts with the term followed by space or punctuation.	English: against, against, aganist, decline, declined, denied, deny, disagree, disagreed, disapporve, disapprove, disapproved, dissaprove, dissent, do not, do not accept, do not approve, do not proceed English: do not support, do nothing, don't, don't accept, don't agree, don't approve, don't support, dont agree, dont support, dont't proceed, i deny, i disagree, i dissent, i do not accept, i do not agree, i do not approve, i do not support, i don't accept English: i don't agree, i don't approve, i don't support, i dont agree, i dont support, i refuse, i reject, ino, nae, nah, nay, negative, negative vote, no, no support, nope, not accept, not agree English: not approve, not approved, not in favor, not support, oppose, opposed, refuse, refused, reject, rejected Non-English: contra, no apruebo, non, non approvo, não, нет, 不会, 不可以, 不同意, 不支持, 不是, 不能, 反对, 反对, 否, 没有

Table 8: Reject stance heuristic registry (continued)

Tier	Rule	Terms
Suffix	Right-trimmed choice ends with a space plus the term.	English: against, against, aganist, decline, declined, denied, deny, disagree, disagreed, disapporve, disapprove, disapproved, dissaprove, dissent, do not, do not accept, do not approve, do not proceed English: do not support, do nothing, don't, don't accept, don't agree, don't approve, don't support, dont agree, dont support, dont't proceed, i deny, i disagree, i dissent, i do not accept, i do not agree, i do not approve, i do not support, i don't accept English: i don't agree, i don't approve, i don't support, i dont agree, i dont support, i refuse, i reject, ino, nae, nah, nay, negative, negative vote, no, no support, nope, not accept, not agree English: not approve, not approved, not in favor, not support, oppose, opposed, refuse, refused, reject, rejected Non-English: contra, no apruebo, non, non approvo, não, нет, 不会, 不可以, 不同意, 不支持, 不是, 不能, 反对, 反对, 否, 没有
High-confidence contains	Choice contains a space-delimited or punctuation-delimited registry term. Confident matches resolve to the stance unless a higher-priority tier already matched.	English: against, against, aganist, decline, declined, denied, deny, disagree, disagreed, disapporve, disapprove, disapproved, dissaprove, dissent, do not approve, do not support, do nothing, don't accept English: don't agree, don't approve, don't support, dont agree, dont support, dont't proceed, i deny, i disagree, i dissent, i do not agree, i do not approve, i do not support, i don't agree, i don't approve, i don't support, i refuse, i reject, ino English: nae, nah, nay, negative vote, no support, nope, not accept, not agree, not approve, not in favor, not proceed, not support, oppose, opposed, refuse, refused, reject, rejected Non-English: contra, non approvo, não, нет, 不会, 不可以, 不同意, 不支持, 不是, 不能, 反对, 反对, 否, 没有

Table 8: Reject stance heuristic registry (continued)

Tier	Rule	Terms
Low-confidence contains	Low-confidence-only matches resolve to Other; conflicts remain visible in the audit trail.	None

Table 9: Approve stance heuristic registry

Tier	Rule	Terms
Exact	Trimmed lower-case choice equals the canonical term.	English: accept, affirmative, agree, agreed, approve, approve!, approved, aye, for, i agree, i support, in favor, ok, pass, positive, positive vote, proceed, s English: support, syes, y, yae, yay, yes, yes! Non-English: a favor, approvo, aprovar, apruebo, oui, si, sim, sí, да, 会, 可以, 同意, 支持, 是, 有, 能, 贊成, 赞成
Prefix	Left-trimmed choice starts with the term followed by space or punctuation.	English: accept, affirmative, agree, agreed, approve, approve!, approved, aye, for, i agree, i support, in favor, pass, positive, positive vote, proceed, support, syes English: yae, yay, yes, yes! Non-English: a favor, approvo, aprovar, apruebo, oui, si, sim, sí, да, 可以, 同意, 支持, 贊成, 赞成
Suffix	Right-trimmed choice ends with a space plus the term.	English: affirmative, agreed, approve!, approved, aye, for, i agree, i support, in favor, pass, positive, positive vote, proceed, syes, yae, yay, yes, yes! Non-English: a favor, approvo, aprovar, apruebo, oui, si, sim, sí, да, 同意, 支持, 贊成, 赞成
High-confidence contains	Choice contains a space-delimited or punctuation-delimited registry term. Confident matches resolve to the stance unless a higher-priority tier already matched.	English: affirmative, agreed, approve!, approved, aye, in favor, positive vote, syes, yae, yay, yes! Non-English: a favor, approvo, aprovar, apruebo, oui, sim, sí, да, 同意, 支持, 贊成, 赞成

Table 9: Approve stance heuristic registry (continued)

Tier	Rule	Terms
Low-confidence contains	Low-confidence-only matches resolve to Other; conflicts remain visible in the audit trail.	English: accept, agree, approve, support, yes

Table 10: Abstain stance heuristic registry

Tier	Rule	Terms
Exact	Trimmed lower-case choice equals the canonical term.	English: abstain, i abstain, maybe, not sure
		Non-English: abstenção, abster, je m'abstiens, 弃权, 棄権
Prefix	Left-trimmed choice starts with the term followed by space or punctuation.	English: abstain, i abstain, maybe, not sure
		Non-English: abstenção, abster, je m'abstiens, 弃权, 棄権
Suffix	Right-trimmed choice ends with a space plus the term.	English: abstain, i abstain, maybe, not sure
		Non-English: abstenção, abster, je m'abstiens, 弃权, 棄権
High-confidence contains	Choice contains a space-delimited or punctuation-delimited registry term. Confident matches resolve to the stance unless a higher-priority tier already matched.	English: abstain
		Non-English: abstenção, abster, je m'abstiens, 弃权, 棄権
Low-confidence contains	Low-confidence-only matches resolve to Other; conflicts remain visible in the audit trail.	None

The raw classifier stores the resolved stance, match source, matched pattern, canonical pattern, all match sources, and a conflict indicator. Confident exact, prefix, suffix, and high-confidence contains matches determine the raw stance. Low-confidence contains terms are still stored in the

Table 11: Audit of one-sided stance recoding for the approval-stance sensitivity analysis

When	Proposal pattern	Proposals	Rows	Spaces
Before recoding	Reject, no Approve	4,274	14,295	1,313
Before recoding	Approve, no Reject	1,945	7,708	892
After recoding	No Approve or Reject labels	75,408	273,825	15,370
After recoding	Approve and Reject both present	51,873	124,393	7,378

Notes: Before recoding rows count undocumented one-sided proposals that were normalized to Other. After recoding rows count the final post-recoding buckets used by the stance sensitivity audit.

audit trail, but by themselves resolve to "Other"; when they occur alongside a confident match, the confident match determines the stance. Cases containing both Approve and Reject evidence are not silently discarded at this stage: the raw classifier resolves by tier order and records an approve/reject conflict flag for later review.

Targeted contextual checks and final audit. After raw registry classification, the notebook writes raw audit CSVs for choice-level matches and proposal-level stance states. It then applies the notebook-level targeted checks and writes a second pair of audit CSVs for the post-notebook-rule checkpoint. Finally, documented metadata rules apply whole-space exclusions, whole-proposal overrides, per-proposal stance-sequence overrides, and one-sided normalization. The final checkpoint writes choice- and proposal-level audit CSVs after overrides. Targeted checks included choices beginning with English negation tokens "dont", "do not", and "don't"; "not" was evaluated but excluded because it was too inclusive. These tokens proxy for proposal-rejecting choices that are often paired with approval choices that pattern matching alone cannot identify (e.g., Buy vs Don't buy or Transfer vs Don't transfer). Stance-specific targeted terms included, for Reject, "neither" and "none of the above"; for Abstain, "Invalid question/options", "revisit", "rework", "revision", and "discuss". Terms evaluated for Approve, including "option", "price", "$", and "will", were not sufficiently associated with approval choices under our definition. Manual inspection also covered stance-bearing emojis.

C.2 Large Language Models (LLMs) Evaluations

To assess the robustness of our choice stance classification, we compare the heuristics with external LLM evaluations of both open source and commercial models, namely Llama 3.1 and 3.3 70 Billion Instruct, OpenAI GPT 4o API (OAI), and ChatGPT 4o. Because LLMs may themselves be sensitive to presentation order, we do not use them as behavioral subjects or as direct evidence of voter bias; instead, we use them only as auxiliary stance classifiers and validate their labels against rule-based heuristics, multiple model outputs, and human annotations. The prompts for the LLMs are reported in Appendix C.2. Llama 3.3 was run only on non-Basic proposals, so its labels cover a sparser subset of the proposal universe than the other LLMs.

Stance detection. We instructed each LLM to assign a stance (as in 'Abstain', 'Approve', 'Other', or 'Reject') to each choice in a proposal, and to assign 'Other' to all choices of proposals not being true governance proposals, to follow closely our heuristics classification goals. The use of LLMs also allows us to examine whether the body of a proposal implicitly or explicitly contains a cue pointing readers toward a voting direction. We treat this *suggested stance* as an LLM-coded textual cue in the proposal text, not as verified author intent or evidence that voters were persuaded. We thus obtain two variables for each LLM that respectively capture whether and in which direction the text suggests a vote, and classify each choice as 'Abstain', 'Approve', 'Other', or 'Reject'.

We note that some proposals could not be parsed correctly by the LLMs, and in a few instances, a different number of choices was interpolated by the LLM compared to the original proposal. These errors are typically associated with proposals having a large number of choices (> 10). We discarded the proposals where we observed this behavior. With this approach, we are able to classify a subset of choices for each LLM (Llama: 260 187 choices, corresponding to 76.50% of all choices; OAI: 278 356, i.e. 81.84%; ChatGPT: 171 766, i.e. 50.50%).

To measure the agreement across our classification methods, we compute Cohen's Kappa between our heuristics and each LLM, using model-specific denominators so Llama 3.3's sparser coverage is handled consistently. The values are 0.69 for Llama 3.1, 0.68 for Llama 3.3, 0.36 for OAI, and 0.61 for ChatGPT, indicating substantial, substantial, fair, and substantial agreement, respectively. The agreement across our classification methods is also displayed in Fig. 7, highlighting two patterns: (i) Llama 3.1 has the highest agreement with the heuristics, and (ii) the proposals we removed due to classification problems from LLMs are mostly from the 'Other' stance, which is often associated with proposals with a large number of choices in the heuristics classification (see Fig. 9).

Suggested stance detection. We also asked whether the proposal text itself appeared to steer voters toward a particular outcome. For each proposal, an LLM would identify a suggested voting direction, such as Approve, Reject, Abstain, or Other, or it would say that the text did not suggest any voting direction. We treated the latter as a reviewed but neutral proposal, not as missing data. A proposal was considered unreviewed and discarded only when a given source produced no usable label after parsing and coverage filters.

The share of proposals with no detected voting suggestion ranges from 53.3% for ChatGPT to 72.5% for Llama 3.3. Specifically, no-suggestion labels account for 57.9% of reviewed, nonconflict proposal-level outputs for Llama 3.1, 72.5% for Llama 3.3, 64% for OAI, and 53.3% for ChatGPT. These shares are source-specific: each model is measured only on proposals for which it produced a usable, nonconflicting label. This distinction is especially important for Llama 3.3, which covers a smaller sample of proposals: 89 479 reviewed proposals, 37 859 not-reviewed proposals, 8998 Basic-proposal coverage exclusions, 67 conflicts, 64 831 no-suggestion proposals, and 24 581 positive suggested-stance proposals. Thus, a missing Llama 3.3 label should not be read as a neutral pro-

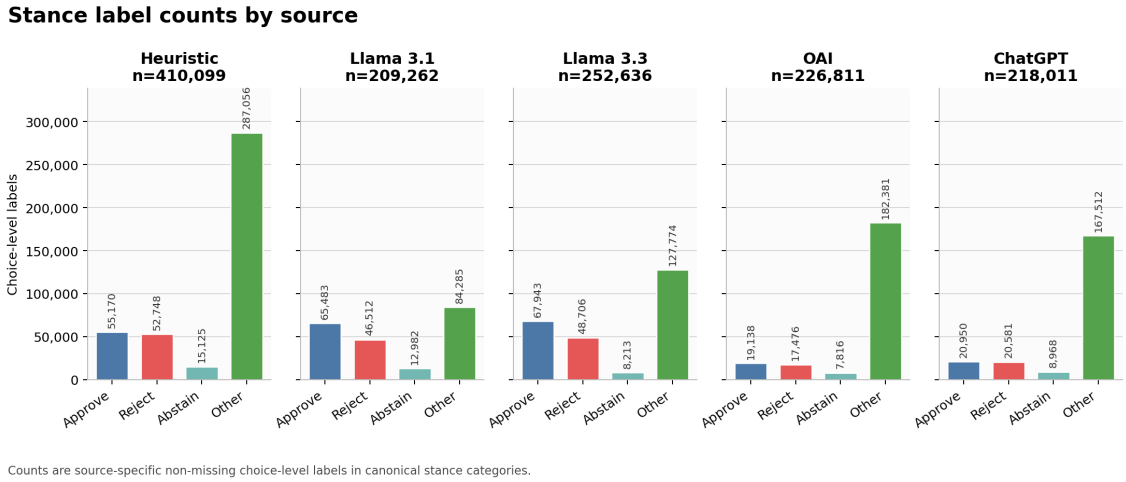


Fig. 7: **Agreement across stance classification methods**. Bars count source-specific non-missing choice-level stance labels for the heuristic, Llama 3.1, Llama 3.3, OAI, and ChatGPT sources. Labels are counted within each source-specific denominator and are limited to Approve, Reject, Abstain, and Other.

posal or as a no-suggestion decision. When an automated model does detect a positive suggested stance, the stance is almost always Approve; Reject, Abstain, and Other cues are rare (see Fig. 8).

Finally, we ask whether a detected suggestion points to the stance that later received the most voting power. In Table 12 we summarize how model-detected textual cues align with realized voting outcomes, using each model's own reviewed sample. Approve is the only positive suggested-stance category with broad coverage, accounting for at least 99.8% of positive automated-model predictions. Across automated models, Approve precision ranges from 0.64 to 0.77, while recall ranges from 0.46 to 0.58, and F1 ranges from 0.54 to 0.65. Reject, Abstain, and Other positive suggested-stance categories appear in too few cases for stable interpretation, so their metric cells should be read as sparse diagnostics rather than as source-level performance estimates.

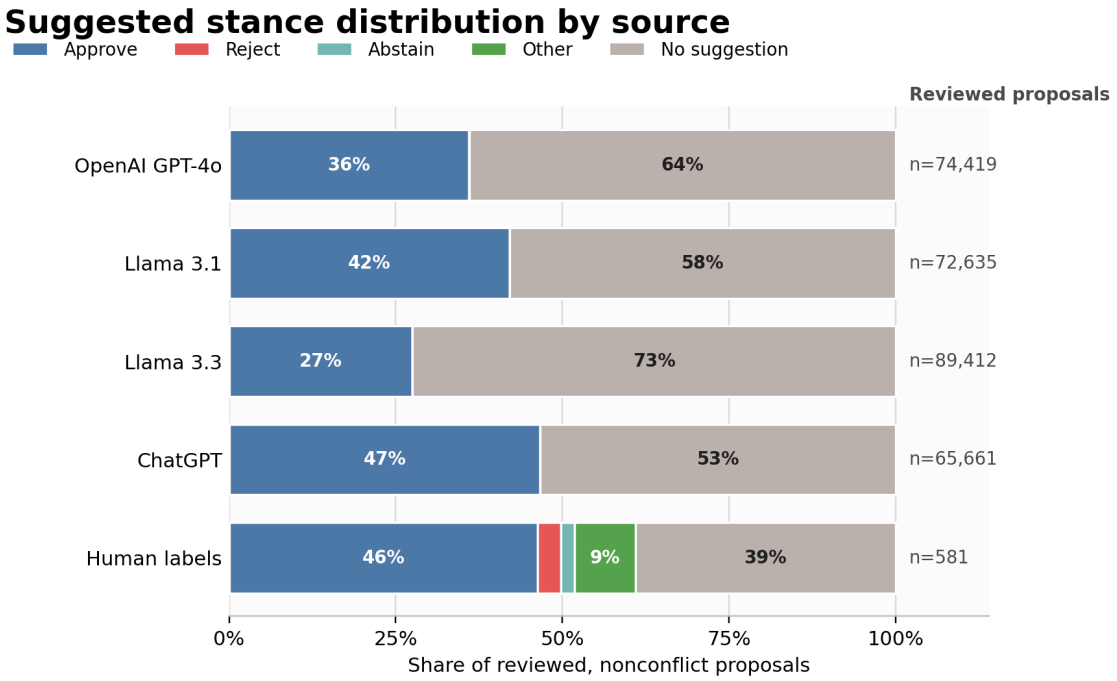


Fig. 8: **Suggested-stance labels by source.** Bars show the share of proposals for which the source found no suggested voting direction or found a textual cue suggesting Approve, Reject, Abstain, or Other. Details about human labels are available in Sec. C.3, but anticipate here the relevant result.

Table 12: Detected suggested stance compared with the winning stance

Source	Stance	Pred.	Winner	TP	Precision	Recall	F1
OpenAI GPT-4o	APPROVE	26737	40070	20509	0.767064	0.511829	0.613978
OpenAI GPT-4o	REJECT	31	5376	11	0.354839	0.002046	0.004069
OpenAI GPT-4o	ABSTAIN	1	749	0	Sparse ($n < 30$)		
OpenAI GPT-4o	OTHER	10	28017	7	Sparse ($n < 30$)		
Llama 3.1	APPROVE	30540	39422	22676	0.742502	0.575212	0.648238
Llama 3.1	REJECT	3	5285	1	Sparse ($n < 30$)		
Llama 3.1	ABSTAIN	0	740	0	Sparse ($n < 30$)		
Llama 3.1	OTHER	0	26986	0	Sparse ($n < 30$)		
Llama 3.3	APPROVE	24523	34870	16016	0.653101	0.459306	0.539323
Llama 3.3	REJECT	23	4884	6	Sparse ($n < 30$)		
Llama 3.3	ABSTAIN	0	466	0	Sparse ($n < 30$)		
Llama 3.3	OTHER	1	49000	1	Sparse ($n < 30$)		
ChatGPT	APPROVE	30599	34640	19477	0.636524	0.562269	0.597097
ChatGPT	REJECT	0	4558	0	Sparse ($n < 30$)		
ChatGPT	ABSTAIN	0	701	0	Sparse ($n < 30$)		
ChatGPT	OTHER	0	25578	0	Sparse ($n < 30$)		

Notes: For each source, we treat the detected suggested stance as a prediction (column Pred.) and we count how many times the prediction was actually true (column TP) over all the time stance a given stance won (column Winner). With this framework, we can compute precision, recall, and F1 to capture alignment between textual suggestions and realized voting outcomes.

Prompts for LLMs

Llama models

```
Reading the title, body, and choices of following proposal, you have to assign one of
the following category to each choice in the proposal:

- APPROVE: supports implementing the proposal's main action or change. This includes
adopting new policies, updating code, adding members, transferring money, establishing
partnerships, organizing events, or any action that moves the proposal forward (including
reversals or cancellations if that's what the proposal is asking for).
- REJECT: opposes implementing the proposal's main action - it wants to maintain the
current state and prevent the proposed change from happening.
- ABSTAIN: leaves the decision to other voters.
- OTHER: if the choice cannot be categorized as either APPROVE, REJECT, or ABSTAIN

A proposal can have multiple choices with the same category.

You must label ALL choices of a proposal as OTHER when:

- There is only one choice in a proposal.
- The choices are opinions on issues unrelated to the DAO (e.g., opinion polls about politics,
weather, life-style, or coins).
- The set of choices of a proposal contains one or more APPROVE choices, but no REJECT choice
(and vice versa).

Please do **not** repeat the question in your answer, and include _only_ one valid JSON
object containing the following properties:

*   effect: Briefly describe the primary consequence(s) of the proposal if it is approved.
Use N/A if the proposal would not result in any direct action or immediate consequence.
*   stances: The stance of all the choices in a JSON array; i.e., if there are 3 choices, the
output should look like [x,y,z], where x, y, z are the stance of the first, second and
third choice.
*   suggested\_stance: If the way the proposal is written (considering title, body, and
choices) in a way that suggests or invites voters to vote for choices with a particular
stance, please write it here, else write N/A.


Do not add any other text to your answer.

Here are the title, body, and choices (separated by '{sep}'), of the proposal:

---

### Proposal:

**Title:**
{title}

**Body:**
{body}

**Choices:**
{choices}
```

OpenAI

```
Your task is to categorize the stance of the choices of proposals from
Decentralised Autonomous Organisation (DAO) on Snapshot.org.

It is crucial to be precise.

Every choice can have one of the following stances:

- APPROVE: selecting this choice brings about a change in the DAO, for instance a change in the
code or in the members of the board or in the governance more broadly, or it involves
transferring money, or establishing a new partnership, or organizing or actively participating
in an event.
- REJECT: selecting this choice cancels the changes caused by choices with APPROVE stance.
- ABSTAIN: selecting this choice leaves the decision to other voters.
- OTHER: if the choice cannot be categorized as either APPROVE, REJECT, or ABSTAIN

You must labels _all_ choices of a proposal as OTHER when:

- There is only one choice in a proposal.
- The choices are opinions on issues unrelated to the DAO (e.g., opinion polls about politics,
weather, life-style, or coins).
- The set of choices of a proposal contains one or more APPROVE choices, but no REJECT choice
(and vice versa).

Please do **not** repeat the question in your answer, and include _only_ one valid JSON object containing
the following properties:

*   effect: Briefly describe the primary consequence(s) of the proposal if it is approved. Use N/A
if the proposal would not result in any direct action or immediate consequence.
*   stances: The stance of all the choices in a JSON array; i.e., if there are 3 choices, the output
should look like [x,y,z], where x, y, z are the stance of the first, second and third choice.
*   suggested\_stance: If the way the proposal is written (considering title, body, and choices) in a way
that suggests or invites voters to vote for choices with a particular stance, please write it here,
else write N/A.

Do not add any other text to your answer.
```

C.3 Human Evaluations

We perform an additional robustness check on the heuristic stance classification. Our aim is to measure the extent to which our heuristics agree with human evaluations, thereby assessing their external validity. We first create a stratified multi-language sample of random proposals and then assign these to crowdworkers for evaluation; the evaluation was implemented with NodeGame[5].

Sample construction We select a set of representative proposals as follows. First, we compute the embeddings of the content of each proposal (title, body, and choices) using a multilingual sentence transformer model (intfloat/multilingual-e5-large); second, we reduced the dimensionality using UMAP; third, we computed 50 coherent clusters using K-Means; finally, we picked 20 random proposals within each cluster for a total of 1,000 proposals to further sample in the next step.

Next, we remove all proposals with less than two choices and more than 10 choices (more than 10 choices would be tiresome for humans to review); then, to ensure multi-language support, the evaluated sample includes proposals in the most common non-English languages in the dataset: Chinese (20), Japanese (5), and Spanish (5). The remainder (560) was filled exclusively from English-language entries[8]. To ensure diversity across different content clusters, we sampled roughly equally from each cluster.[9]

Crowd evaluation Using the online labor market for research Prolific, we assign a group of evaluators (N = 375) the task of manually labeling five proposals, so that each proposal is reviewed at least 3 times. We used platform filters to select crowdworkers fluent in English and, in addition, having Japanese, Chinese, or Spanish as primary language for the non-English sets. Because crowdworkers do not likely have background knowledge in DAOs, we implemented a two-step quiz procedure. At step 1., we explained what DAOs are and how voting takes place in DAOs (one multiple-choice quiz question). At step 2. we explained the rules for labeling the choices of governance proposals (five mandatory multiple-choice quiz questions). Participants who failed twice at any step were required to leave the task; this ensured an informed base for the decisions of all crowdworkers.

For each proposal, crowdworkers had three subtasks: (i) to label the stance of each choice ('Abstain', 'Approve', 'Other', or 'Reject'), (ii) to mark if the proposal was real or spam, and (iii) to indicate whether the text of the proposal is written in a way that suggests a given action (e.g., 'Approve').

At the end of the task, crowdworkers were asked to express how clear the task and proposals were. 86.1% of them completely or somehow agreed that the task was clear, while 74.9% of them completely or somehow agreed that the proposals were clear. Participants who completed the entire

[8] This is to avoid assigning proposals of different languages to evaluators.

[9] Korean proposals were selected but eventually not evaluated due to an assignment bug, so the total number of evaluated proposals is 590.

process received a fixed payment of US$2 for the task, and it took on average 14.09 minutes to complete it. Further information on crowdworkers' demographics is reported in Table 13.

Comparison of heuristics and human evaluations In the final step, we aggregate the evaluations provided by different crowdworkers and assign each stance a unique label ('Abstain', 'Approve', 'Other', or 'Reject').

To do so, we adopt a majority rule approach and assign a label to a stance only when it is possible to identify one single most occurring label; if two or more evaluations are tied as the most occurring, we leave the label unassigned. For instance, if the evaluations for one stance are [Reject, Reject, Other, Other], that stance is left unassigned. Next, we fill in unassigned stances when the label can be inferred from the remaining choices of the same proposal. Specifically, proposals that include one 'Approve' typically also include a 'Reject' option (and vice-versa). We thus identify cases where a proposal has exactly one 'Approve' ('Reject') choice, one unassigned choice, and all remaining choices labeled as 'Other' or 'Abstain'. If at least one annotator assigned the missing label to the 'Reject' ('Approve') option, then we label the non-assigned choice accordingly. As an illustrative example, a proposal with choices labeled as {N/A, Reject, Abstain} after the first step will be updated to {Approve, Reject, Abstain} if at least one crowdworker evaluated the missing label as 'Approve'.

We flag evaluations that are potentially carried out with low effort in two complementary ways: first, we identify those where users spent unusually little time completing their task (i.e., less than five minutes, considering that the task requires approximately ten minutes to be conducted). Second, we group proposals per language, normalize their body lengths, and identify those with short evaluation time (less than 20 seconds) despite being relatively long proposals (above the 33rd percentile of all proposals).

Lastly, we measure to what extent the heuristics are consistent with the crowdworkers' evaluations. The Cohen's Kappa of the heuristics and the human evaluators is 0.71, which indicates substantial agreement. This comparison covers 1603 choices, and the result does not change when removing the 6 evaluators flagged as low effort.

Table 13: Demographic breakdown of evaluators

Category	All evaluators (N =354)	Fast evaluators (N = 6)	Normal evaluators (N = 348)
Gender			
Male	210	2	208
Female	142	4	138
Non-binary	2	0	2
Age Group			
18-20	11	0	11
21-25	43	1	42
26-30	66	3	63
31-35	48	2	46
36-40	59	0	59
41-45	35	0	35
46-50	23	0	23
51-55	22	0	22
56-60	18	0	18
61-65	14	0	14
66-70	7	0	7
71-75	8	0	8
Education			
High-School	65	0	65
College	147	3	144
Grad School	142	3	139
Prior knowledge of DAOs			
No	286	6	280
Yes	68	0	68
Time to conduct tasks			
Avg. Time	14.09	3.98	14.27
Task clarity			
Completely agree	165	1	164
Somewhat agree	140	4	136
Neither agree nor disagree	22	1	21
Somewhat disagree	26	0	26
Completely disagree	1	0	1
Proposal clarity			
Completely agree	118	0	118
Somewhat agree	147	4	143
Neither agree nor disagree	39	0	39
Somewhat disagree	45	2	43
Completely disagree	5	0	5

Notes: the evaluators are N = 375. The demographics are computed on the restricted set of users who completed the last task (i.e., providing user information and feedback).

D Additional Material and Info

In this section, we report additional regression and statistical tables to support the claims of the paper, divided by the section of the main paper in which they are most closely referenced.

D.1 Position Bias

Table 14 and Table 15 show descriptive statistics about voting power, participation, and the three biases under study.

Table 14: Choice position, voting-power share, and individual-vote share

	Voting Power		Number of votes		Percent	N
Choice Pos.	Mean	SD	Mean	SD		
1	0.70	0.39	0.71	0.37	31.96	127,281
2	0.22	0.35	0.22	0.32	31.96	127,281
3	0.14	0.26	0.13	0.24	9.87	39,303
4	0.14	0.24	0.13	0.22	3.48	13,839
5	0.11	0.22	0.11	0.19	2.04	8,130
6	0.09	0.18	0.08	0.15	1.28	5,110
7	0.07	0.15	0.07	0.13	0.97	3,853
8	0.06	0.14	0.06	0.13	0.81	3,209
9	0.05	0.13	0.06	0.12	0.69	2,766
10	0.05	0.12	0.05	0.10	0.62	2,482
10+	0.01	0.06	0.01	0.05	16.31	64,964

Notes: The sample is the final cleaned dataset with 398 218 choices. Outcomes are mean choice-level voting-power share and mean favorite-choice vote share by choice position; N indicates the number of choices with a given position index; positions above 10 grouped as 10+.

Table 15: Voting power, frequency of approval and author choices of the first choice by voting type

	Voting Power		Approve		Author Choice		All	
type	Mean	SD	Mean	SD	Mean	SD	Percent	N
Basic	0.79	0.34	0.77	0.42	0.84	0.37	7.05	8,975
Single-choice	0.71	0.39	0.38	0.49	0.73	0.44	85.95	109,398
Approval	0.42	0.37	0.20	0.40	0.67	0.47	1.21	1,537
Quadratic	0.50	0.42	0.30	0.46	0.69	0.46	1.41	1,793
Ranked-choice	0.40	0.38	0.19	0.39	0.51	0.50	0.71	907
Weighted	0.47	0.41	0.36	0.48	0.64	0.48	3.67	4,671

Notes: the sample is the final cleaned dataset with 398 218 choices.

D.2 Approval Bias

Table 16 reports the stance (Approve, Reject, Abstain, Other) shares for every position across all proposals. Fig. 9 reports the stance counts by the number of choices in the proposal. Together, they show that the stance Other is especially common in proposals with many choices, while two- and three-choice proposals more often follow an Approve/Reject/Abstain structure.

Fig. 10 tackles a different question. Panel A compares the average voting-power share of choices with each stance. Panel B reports, within each stance and choice-count group, the share of choices that appear in the first ballot position. Thus, Fig. 9 describes how common each stance is, whereas Fig. 10 describes voting power and ballot placement conditional on stance. From Panel B we learn that when an Approve choice is found in a proposal this is almost placed in position 1 for proposals with two or three choices; overall Approve is the most the common stance across all sub-panels.

Finally, Fig. 11 extends Fig. 3 in the maintext, disaggregating the excess/deficit of voting power by the number of choices in a proposal. The qualitative findings remain the same.

Table 16: Stance distribution by choice position

	Approve		Reject		Abstain		Other		All	
Choice Position	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Percent	N
1	0.40	0.49	0.00	0.07	0.00	0.01	0.59	0.49	31.96	127,281
2	0.01	0.12	0.39	0.49	0.00	0.05	0.59	0.49	31.96	127,281
3	0.02	0.15	0.03	0.18	0.35	0.48	0.60	0.49	9.87	39,303
4	0.03	0.16	0.06	0.24	0.03	0.17	0.89	0.32	3.48	13,839
5	0.02	0.14	0.02	0.13	0.06	0.24	0.90	0.30	2.04	8,130
6	0.01	0.11	0.01	0.12	0.01	0.11	0.96	0.19	1.28	5,110
7	0.02	0.13	0.01	0.08	0.00	0.07	0.97	0.16	0.97	3,853
8	0.01	0.10	0.01	0.11	0.00	0.06	0.98	0.16	0.81	3,209
9	0.01	0.11	0.00	0.06	0.00	0.04	0.98	0.13	0.69	2,766
10	0.01	0.10	0.01	0.08	0.01	0.11	0.97	0.16	0.62	2,482
10+	0.01	0.09	0.00	0.04	0.00	0.02	0.99	0.10	16.31	64,964

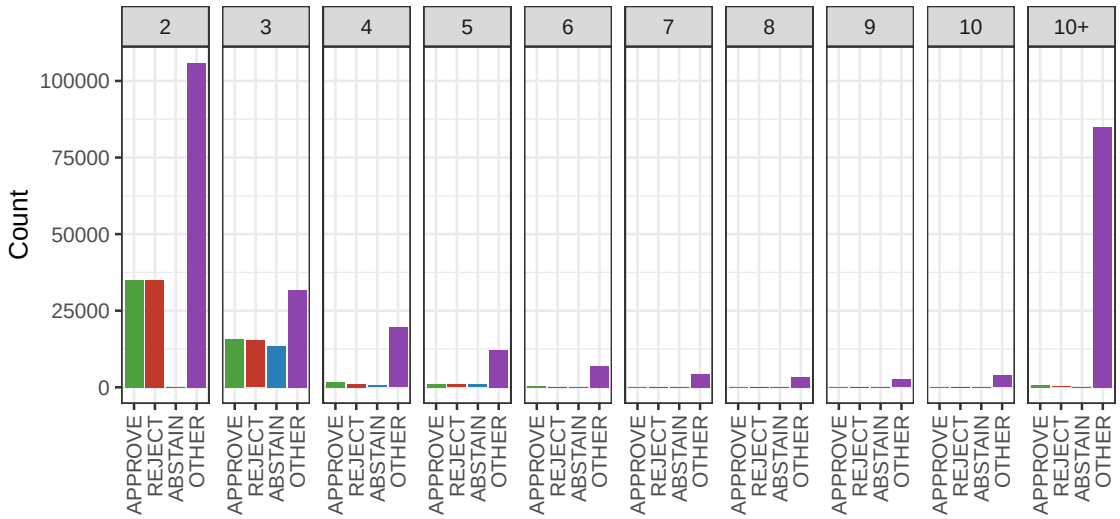


Fig. 9: **Choice stance counts by number of choices in a proposal (full sample).**

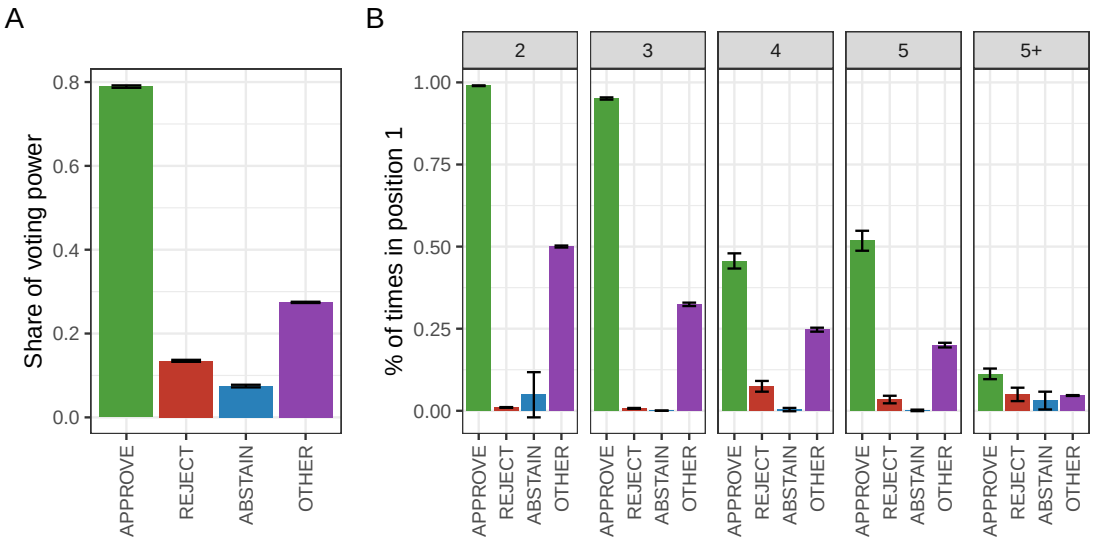


Fig. 10: **Voting power and position by stance (full sample). A.** Mean choice-level voting-power share by stance. **B.** Share of choices with a given stance that are in position 1, by number of choices in the proposal. These shares are conditional on stance and therefore do not describe the overall distribution of stances. Confidence intervals are 95% confidence intervals of the means.

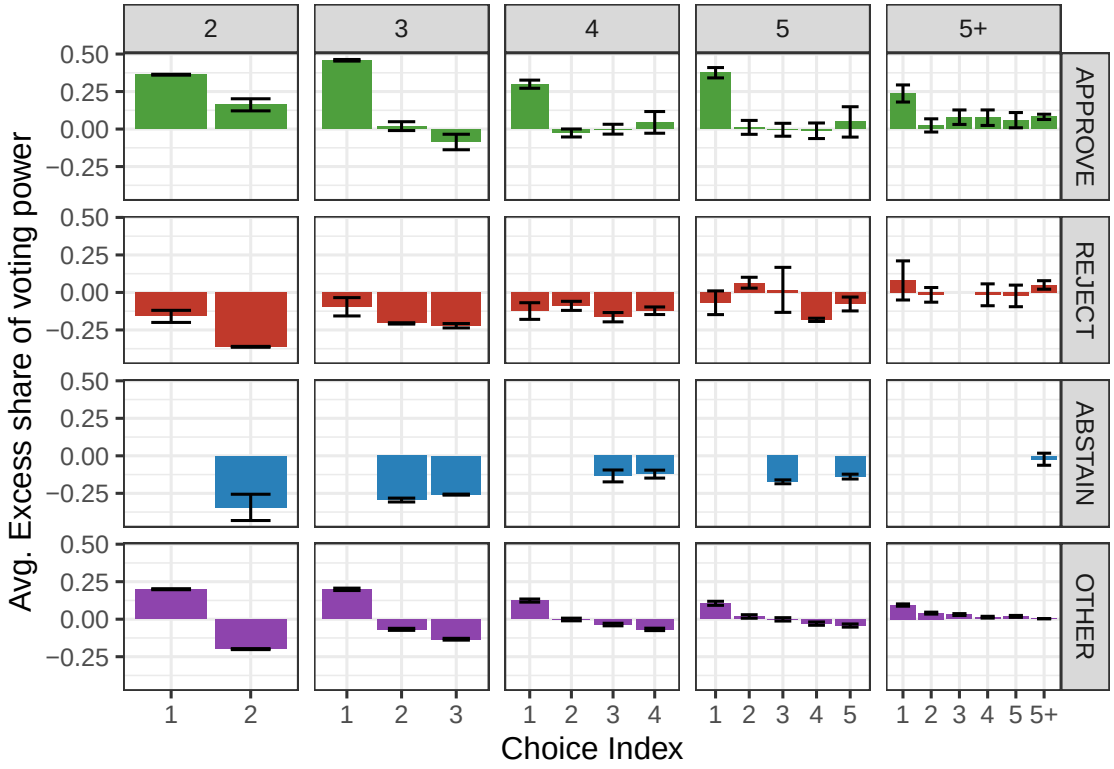


Fig. 11: **Excess voting power by stance, choice position, and number of choices (full sample).** Only stance-position-num-choice combinations with more than 10 observations are kept in the plot. Error bars are 95% confidence intervals of the means.

Author Bias Fig. 12 shows the boost in excess voting power when an author chooses a choice at a given position with a given stance.

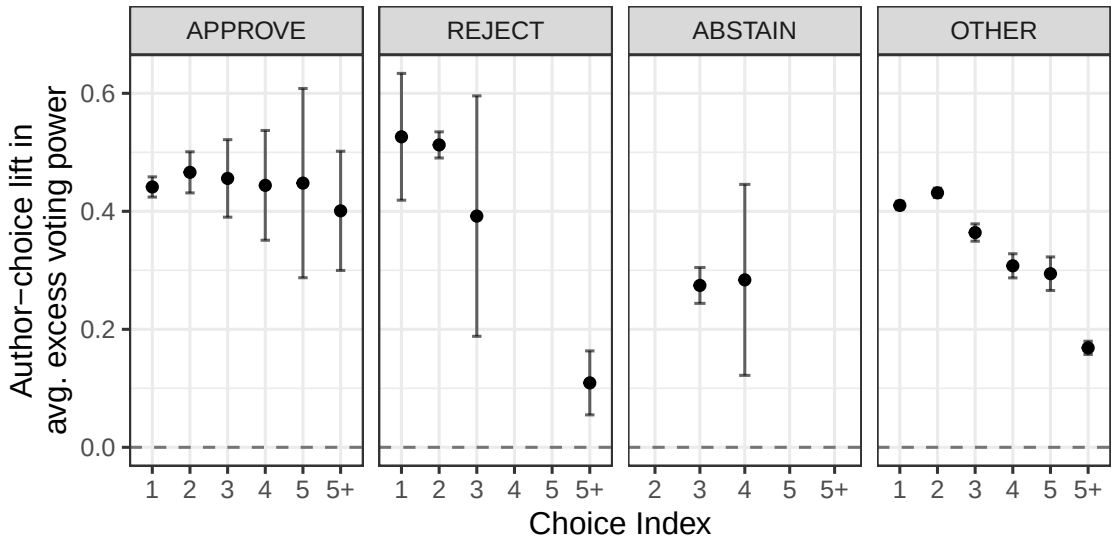


Fig. 12: **Author boost by stance and choice position.** Each dot in the plot is the difference between the point estimate of the average excess voting power of a choice of an author vs other choices. Error bars are 95% confidence intervals of the means.

D.3 Regression Robustness Specifications and Sensitivity Tables

Table 17 reports the shared robustness grid for the three focal AMEs. Cells show response-scale AMEs with Bonferroni/FWER-adjusted DAO-space clustered-bootstrap confidence intervals, using the same 36-AME family as Figure 5. This replaces separate one-off inference tables for the primary, author-voted-only, and participation-control rows, so every shared robustness specification is reported with the same three-bias contrast set and uncertainty convention. We include *Vote-count* as a shared outcome robustness row, keeping the vote-count-share sensitivity inside this common grid rather than in a standalone table. The first-vs-second position AME is a counterfactual prediction contrast, not a first-vs-rest effect and not the coefficient on choice rank. For each fitted model, we predict outcomes on the estimation sample with choice rank set to 1 and then to 2, recomputing the squared-rank term in each counterfactual, and average the response-scale difference.

First-choice position sensitivity Tables 18 and 19 report the sensitivity check that replaces the rank-plus-rank-squared position controls with a binary first-choice indicator. The coefficient ladder shows the main single-preference specification, while the AME grid mirrors the shared robustness specifications.

Table 20 documents the DAO spaces excluded in the *No top-5* row of the shared robustness table.

Table 17: Three-bias AMEs across shared robustness specifications

Specification	Author choice	First vs second	Approve stance
Main	58.78 [53.82, 63.74]	7.75 [5.41, 10.08]	27.14 [20.52, 33.76]
Vote-count	57.64 [53.46, 61.83]	8.25 [5.71, 10.79]	29.72 [24.25, 35.18]
Auth-voted	65.88 [61.34, 70.42]	4.29 [2.37, 6.21]	7.72 [4.66, 10.78]
Compat full	48.02 [43.34, 52.70]	6.29 [4.24, 8.35]	24.64 [18.65, 30.62]
All systems	44.56 [39.84, 49.27]	1.16 [-0.27, 2.59]	29.07 [22.82, 35.33]
Vote-control	58.78 [53.81, 63.75]	7.74 [5.41, 10.08]	27.14 [20.52, 33.76]
TVL-25	27.12 [7.59, 46.65]	7.65 [-9.97, 25.28]	51.72 [30.48, 72.96]
TVL-50	32.61 [14.46, 50.77]	10.70 [-5.18, 26.57]	46.41 [28.25, 64.56]
TVL-100	32.92 [16.12, 49.73]	11.17 [-4.12, 26.46]	45.13 [27.65, 62.62]
Proposal FE	56.63 [51.09, 62.17]	5.94 [3.75, 8.12]	35.69 [27.25, 44.13]
Low-part.	39.73 [34.70, 44.76]	8.97 [6.05, 11.89]	32.51 [24.59, 40.43]
No top-5	59.31 [55.62, 63.01]	6.91 [5.31, 8.50]	30.08 [25.83, 34.32]

Notes: Cells report response-scale average marginal effects (AMEs) in percentage points with Bonferroni/FWER-adjusted 95% DAO-space clustered-bootstrap confidence intervals in brackets. Intervals are adjusted across the 36 shared AMEs shown in this table. For each specification, the model is fit on the rows selected by that specification. For the author-choice column, the AME is then averaged over the subset of that model's fitted rows where `author_voted`=1. In the *Auth-voted* row, this subset is the whole fitted sample; in the other rows, it is usually only part of the fitted sample. The author-choice column compares author-selected with non-author-selected choices among rows from proposals where the author voted. The first-vs-second column reports the predicted difference between choice rank 1 and choice rank 2 using the rank-plus-rank-squared position specification. The approve-stance column reports the response-scale AME for the approval stance, comparing choices coded as approving with choices coded reject, abstain, or other.

Table 18: First-choice-position sensitivity specifications

	(1)	(2)	(3)	(4)	(5)	(6)
Author choice	3.250***	3.250***	3.237***	3.237***	3.242***	1.937***
	(0.088)	(0.088)	(0.088)	(0.088)	(0.087)	(0.092)
Approving choice	1.301***	1.301***	1.277***	1.277***	1.273***	1.287***
	(0.134)	(0.134)	(0.135)	(0.135)	(0.135)	(0.135)
First choice	1.286***	1.286***	1.259***	1.259***	1.251***	1.257***
	(0.052)	(0.052)	(0.051)	(0.051)	(0.051)	(0.051)
Author voted	-1.499***	-1.499***	-1.494***	-1.472***	-1.474***	-0.883***
	(0.039)	(0.039)	(0.040)	(0.038)	(0.038)	(0.040)
Relative proposal date		-0.000	-0.000	-0.000	-0.000	-0.000
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Number of choices			-0.104***	-0.104***	-0.105***	-0.106***
			(0.016)	(0.016)	(0.016)	(0.016)
Author choice in first six				-0.026*	-0.026*	-0.683***
				(0.011)	(0.011)	(0.044)
Author choice x first-six visibility						1.443***
						(0.092)
Relative choice text length					-0.120+	-0.123+
					(0.064)	(0.063)
Voting-type controls	Yes	Yes	Yes	Yes	Yes	Yes
N	293816	293816	293816	293816	293816	293816

Notes: Coefficient estimates are on the logit scale with DAO-space clustered standard errors in parentheses. Significance markers on coefficients: + p below 0.1, * p below 0.05, ** p below 0.01, *** p below 0.001.

Table 19: First-choice position-control three-bias AMEs across shared robustness specifications

Specification	Author choice	First vs non-first	Approve stance
Main	56.20	19.66	18.57
Vote-count	55.05	20.38	20.95
Auth-voted	64.62	9.10	5.52
Expanded	46.85	19.44	16.67
Full	40.92	17.07	16.03
Vote-control	56.20	19.66	18.57
TVL-25	18.01	43.17	22.27
TVL-50	21.24	45.64	20.03
TVL-100	21.45	45.94	19.32
Proposal FE	54.88	16.37	25.66
Low-part.	34.67	25.68	21.19
No top-5	56.83	18.14	21.63

Notes: Cells report response-scale average marginal effects (AMEs) in percentage points. This grid mirrors the main three-bias summary in Table 17 but replaces the rank-plus-rank-squared position controls with a binary first-choice indicator. The first-vs-non-first column reports the predicted response-scale difference between first-choice and non-first-choice rows.

Table 20: Top spaces excluded in robustness model

Rank	Space
1	cakevote.eth
2	snapshot.dcl.eth
3	pancake
4	index-coop.eth
5	polls.lenster.xyz

D.4 Matching of Top-100 DeFi Protcols

Data collection. We downloaded the complete set of approximately 7,500 protocols returned by DefiLlamas APIs. Protocols classified under centralized-exchange categories were excluded prior to analysis, to focus on decentralized finance protocols only, the kind of protocols one expect to have a DAO on Snapshot.

For each eligible protocol, we retrieved the historical TVL records and filtered to retain only the daily total value locked observations associated with the Ethereum chain. The study period was defined as 1 October 2020 through 30 November 2023, inclusive. Observations outside this window were discarded. Because multiple observations may occur for the same UTC calendar date, records were deduplicated at the UTC-date level, retaining the last available observation for each protocol-date pair.

To ensure sufficient temporal coverage, protocols were required to have Ethereum TVL observations for at least 80% of the study period, corresponding to a minimum of 925 observed days. Protocols below this coverage threshold were excluded. Missing dates were not imputed: no zero-filling, forward-filling, or other interpolation procedure was applied.

Eligible protocols were ranked according to the median of their observed daily Ethereum TVL values during the study period. The median, rather than the mean, was used as the primary ranking statistic to reduce sensitivity to short-lived TVL spikes or extreme observations. Protocols were ordered in descending median Ethereum TVL, with ties resolved alphabetically.

Matching procedures. We constrained matching to the top 100 DefiLlama's protocols ranked as specified above. The fields `governanceID` and `governanceIDs` sometimes contain Snapshot identifiers encoded as `snapshot:<space-id>`; all such identifiers are extracted and stored. Because DefiLlama organises many protocols into parent–child hierarchies, governance identifiers are often attached only to the parent record rather than to each child protocol. We therefore query the website-backing `/lite/protocols2` endpoint to retrieve parent-level governance identifiers, and each child protocol inherits its parent's Snapshot identifiers in a separate parent field, so that the provenance of direct and inherited matches remains distinct. A protocol–space pair is accepted automatically when the Snapshot space appears in either the direct or inherited DefiLlama fields.

The automatic matches are then audited against local Snapshot data. Accepted pairs remain eligible only if the Snapshot space exists in our final cleaned dataset. Going beyond exact matching, we then construct plausible Snapshot identifiers from each protocol's slug and name by normalising text, stripping common version or category suffixes such as `-v2`, and appending common Snapshot suffixes such as `.eth`, `dao.eth`, and `gov.eth`. These fuzzy candidates are used for audit and manual review, but they are not accepted automatically. Protocols that remain unresolved are reviewed manually and the final ressults are shown in Table 21 and 22.

Table 21: Top 100 matched DAO Snapshot-space ranks used in the Top-TVL robustness subset (ranks 1-25)

#	DefiLlama rank	DefiLlama protocol	Snapshot space	Category	TVL	Prop.	Choices
1	2	Lido	lido-snapshot.eth	Liquid Staking	$6,249,613,456	156	329
2	3	Curve DEX	curve.eth	Dexs	$4,584,909,044	124	287
3	4	Convex Finance	cvx.eth	Yield	$4,079,001,323	744	10,238
4	5	Aave V2	aave.eth	Lending	$4,027,204,550	457	1,392
4	35	Aave V1	aave.eth	Lending	$80,000,026	457	1,392
5	6	Compound V2	comp-vote.eth	Lending	$2,697,179,105	5	17
6	7	Uniswap V3	uniswap	Dexs	$2,532,881,366	105	282
6	8	Uniswap V2	uniswap	Dexs	$1,740,993,807	105	282
6	77	Uniswap V1	uniswap	Dexs	$8,548,008	105	282
7	10	Balancer V2	balancer.eth	Dexs	$758,661,665	625	1,561
7	16	Balancer V1	balancer.eth	Dexs	$279,586,764	625	1,561
8	12	Yearn Finance	veyfi.eth	Yield Aggregator	$649,626,401	7	23
9	13	SushiSwap	sushigov.eth	Dexs	$615,839,109	201	585
10	14	Synthetix v1+v2	synthetix-stakers-poll.eth	Synthetics	$592,131,590	2	4
11	15	dYdX V3	dydxgov.eth	Derivatives	$379,380,032	35	74
12	18	Nexus Mutual	community.nexusmutual.eth	Insurance	$267,782,615	56	137
13	20	RenVM	ren-project.eth	Bridge	$244,749,194	31	139
14	21	Frax	frax.eth	Algo Stables	$205,967,755	348	712
15	23	Harvest Finance	harvestfi.eth	Yield Aggregator	$194,147,321	7	15
16	25	Bancor V2.1	bancornetwork.eth	Dexs	$173,900,182	542	1,392
17	26	xDAI Stake Bridge	xdaistake.eth	Canonical Bridge	$168,526,652	214	471
18	27	StakeWise V2	stakewise.eth	Liquid Staking	$135,455,186	73	167
19	28	Badger DAO	badgerdao.eth	Yield Aggregator	$132,500,580	104	280
20	29	Loopring	loopringdao.eth	Dexs	$123,843,786	14	188
21	31	Index Coop	index-coop.eth	Indexes	$101,927,589	977	2,424
21	32	Set Protocol	index-coop.eth	Indexes	$94,881,084	977	2,424
22	36	Stake DAO	stakedao.eth	Yield	$74,601,704	88	259
23	37	Idle	idlefinance.eth	Yield Aggregator	$70,569,685	110	430
24	37	Idle	staking.idlefinance.eth	Yield Aggregator	$70,569,685	101	352
25	39	KEEP Network	keepstakers.eth	Cross Chain Bridge	$70,020,929	10	67

Notes: This part reports Top-TVL sample ranks 1-25. Not all protocols found in DefiLLama can be matched in our Snapshot database; multiple DefiLlama protocols can be matched to the same Snapshot space.

Table 22: Top 100 matched DAO Snapshot-space ranks used in the Top-TVL robustness subset, continued (ranks 26-100)

#	DefiLlama rank	DefiLlama protocol	Snapshot space	Category	TVL	Prop.	Choices
26	39	KEEP Network	threshold.eth	Cross Chain Bridge	$70,020,929	67	214
27	40	Ribbon	gauge.rbn.eth	Options Vault	$61,640,461	35	232
28	40	Ribbon	rbn.eth	Options Vault	$61,640,461	34	88
29	41	Fei Protocol	fei.eth	Algo Stables	$61,030,545	136	423
30	42	Homora V2	alpha-finance-lab.eth	Leveraged Farming	$58,282,004	1	2
31	43	CREAM Lending	cream-finance.eth	Lending	$52,907,346	58	127
31	85	CreamSwap	cream-finance.eth	Dexs	$4,952,526	58	127
32	46	Notional V2	notional.eth	Lending	$43,636,953	30	69
33	47	Origin Dollar	origingov.eth	Yield Aggregator	$40,954,614	65	312
34	47	Origin Dollar	ousdgov.eth	Yield Aggregator	$40,954,614	97	861
35	48	mStable CDP	gov.dhedge.eth	CDP	$40,248,021	70	153
35	88	Chamber Vaults	gov.dhedge.eth	Indexes	$4,386,828	70	153
36	48	mStable CDP	mstablegovernance.eth	CDP	$40,248,021	142	411
36	88	Chamber Vaults	mstablegovernance.eth	Indexes	$4,386,828	142	411
37	49	Saddle Finance	saddlefinance.eth	Dexs	$38,370,975	88	593
38	50	Vesper	vsp.eth	Yield Aggregator	$38,096,025	25	67
39	51	dForce Lending	dforcenet.eth	Lending	$35,203,566	65	131
40	52	Rari Capital	fuse.eth	Yield Aggregator	$33,884,090	144	354
41	53	Unslashed	unslashed.eth	Insurance	$33,712,560	30	79
42	58	SharedStake	sharedstake.eth	Liquid Staking	$29,779,030	50	165
43	59	Rook	rook.eth	Dexs	$28,495,996	43	96
44	60	Ooki	ooki.eth	Lending	$26,325,261	8	16
45	62	Metronome V1	metronome.eth	Yield	$23,430,580	19	57
46	64	Rhino.fi	rhinofi.vote	Bridge	$19,079,766	19	62
47	65	Parallel Protocol V2	mimo.eth	CDP	$18,920,030	114	347
48	66	TrueFi	truefigov.eth	Uncoll. Lending	$17,146,372	77	167
49	67	Bella Protocol	bella	Yield	$15,051,240	1	5
50	69	Pickle	pickle.eth	Yield Aggregator	$13,971,060	51	150
51	71	DFX V2	dfx.eth	Dexs	$13,327,258	32	69
52	72	PoolTogether V3	poolpool.pooltogether.eth	Yield Lottery	$11,460,793	65	151
53	72	PoolTogether V3	pooltogether.eth	Yield Lottery	$11,460,793	24	64
54	74	Gnosis Protocol v1	gnosis.eth	Prediction Market	$9,516,819	82	227
55	75	88mph	88mph.eth	Lending	$9,282,355	11	22
56	76	1inch	1inch.eth	DEX Aggregator	$8,854,452	47	124
57	78	OnX Finance	onx-finance.eth	Yield Aggregator	$8,466,727	12	82
58	79	BarnBridge	barnbridge.eth	Yield Aggregator	$7,990,989	32	75
59	84	Hakka Finance	hakka.eth	Derivatives	$4,957,516	15	50
60	91	PieDAO	piedao.eth	Indexes	$3,850,940	111	273
61	93	Perlin	perlinx.eth	Dexs	$3,814,817	3	6
62	94	Dego Finance	dego	Services	$3,670,970	8	37
63	99	xToken	xxtka.eth	Liquidity Manager	$2,821,323	7	24

Notes: This part reports Top-TVL sample ranks 26-100. Not all protocols found in DefiLLama can be matched in our Snapshot database; multiple DefiLlama protocols can be matched to the same Snapshot space.