

Belief-Aware Pivotal Mechanism for DAO Committees

Nuno Braz 

INESC-ID, Instituto Superior Técnico, Universidade de Lisboa, Portugal

Diogo Poças 

Instituto de Telecomunicações, Instituto Superior Técnico, Universidade de Lisboa, Portugal

Miguel Correia 

INESC-ID, Instituto Superior Técnico, Universidade de Lisboa, Portugal

Abstract

Decentralized Autonomous Organizations (DAOs) increasingly delegate decisions to small committees whose members hold two independent kinds of private information: idiosyncratic preferences over alternatives (what they want) and beliefs about which alternative best serves the organization (what they know). Members have no reason to reveal what they know unless they are incentivized to do so.

Standard voting rules are designed to extract only what members want. If these two pieces of information are not aligned, the organization can end up making a suboptimal decision. Existing mechanisms for eliciting (extracting) beliefs from experts, such as Decision Scoring Rules and Decision Markets, face well-known impossibility results in multi-agent settings under deterministic decision rules, and ignore the fact that agents can have idiosyncratic preferences over alternatives.

This paper proposes a mechanism for binary committee decisions that augments the pivotal mechanism, an instance of the Groves mechanism, with a reward that depends on the outcome distributed after a boolean signal, indicating success or failure of the outcome, is observed. The mechanism aggregates the experts' private information to maximize the probability that its decision agrees with a weighted linear pooling of the experts' beliefs, framing the designer's problem as classification rather than welfare maximization.

An affine family of reward rules is proposed, and the region of parameters that simultaneously satisfy dominant-strategy incentive compatibility, interim individual rationality, and budget feasibility is derived. In informative-belief environments, the mechanism outperforms majority voting, including majority voting augmented with the same outcome-contingent rewards, especially when the committee's average preferences are biased against the superior alternative.

2012 ACM Subject Classification Theory of computation → Algorithmic mechanism design; Applied computing → Electronic commerce

Keywords and phrases Mechanism Design, Decentralized Autonomous Organizations, Committees, Information Aggregation

Digital Object Identifier 10.4230/LIPIcs.AFT.2026.17

Supplementary Material *Software*: <https://github.com/Nunobrazz/committee-simulator>
archived at `swb:1:dir:ad34d49c6e83305835e1e63b3da599a62838a0d1`

Funding This work was financially supported by Project Blockchain.PT – Decentralize Portugal with Blockchain Agenda (Project no 51), WP 6: Digital Assets Management, Call no 02/C05-i01.01/2022, funded by the Portuguese Recovery and Resilience Program (PRR), The Portuguese Republic, and The European Union (EU) under the framework of the Next Generation EU Program. This work was also supported by national funds through Fundação para a Ciência e a Tecnologia, I.P. (FCT) under projects UID/50008/2025 (DOI: <https://doi.org/10.54499/UID/50008/2025>) and UID/50021/2025 (DOI: <https://doi.org/10.54499/UID/50021/2025>).



© Nuno Braz, Diogo Poças, and Miguel Correia;
licensed under Creative Commons License CC-BY 4.0
8th Conference on Advances in Financial Technologies (AFT 2026).

Editors: Aggelos Kiayias and Maria Kyropoulou; Article No. 17; pp. 17:1–17:22



Leibniz International Proceedings in Informatics

Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1 Introduction

A Decentralized Autonomous Organization (DAO) is an organization governed by rules encoded as smart contracts on a blockchain, where members coordinate and make collective decisions without centralized authority. In recent years, DAOs have grown into large-scale entities where thousands of members coordinate pseudonymously [34]. Unlike traditional contracts, which merely articulate rules and can only serve as legal proof, smart contracts can execute the rules they encode. The scale of these organizations creates a significant challenge in decision-making. Several studies [6, 33, 18] confirm that DAOs suffer from low participation rates and slow execution, showing the need for delegated governance structures. In large blockchains such as Bitcoin and Ethereum, the decision-making is done by the maintainers [8, 2]; Polkadot elects a council [30]; project Catalyst has a Community Advisors board and a veteran [31]. Either way, when governance councils are not elected, decision power is concentrated in a small number of members [18]. Empirical and analytical studies confirm that token-based voting is insufficient when participants have heterogeneous stakes and information [18, 14, 36, 11].

Restricting power to a small group or committee amplifies a natural problem: members often have personal interests that are not aligned with the organization's goal. For example, a committee governing a lending DAO might vote to approve a highly risky token as collateral simply because they personally own large amounts of it, exposing users to severe financial risk. Mechanisms such as majority voting and quadratic voting [19, 4] are designed to account only for what members *want* and not for what they *know* about which alternative benefits the organization. Furthermore, any non-continuous voting mechanism necessarily loses information by restricting experts to discrete reports. Talmon [35] identifies DAOs as an important application domain for Multi-Agent Systems and calls for the community to address their governance challenges. Our work addresses this by proposing a decision-making mechanism with formal incentive guarantees that improves collective decisions through better information aggregation.

Organizations often do an evaluation of the consequences of their decisions. These evaluations are usually informal: performance reviews, retrospectives, reputation effects. In a DAO, the evaluation can be formalized: the community can express whether a decision led to a positive or negative outcome after a predetermined period. More quantitative criteria also serve, such as treasury growth, total value locked, user adoption. The evaluation task is simpler than the original decision, as it concerns observed outcomes rather than predictions. Because smart contracts can hold funds in escrow and release them conditionally, members' compensation can naturally be tied to the actual consequences of the decisions they supported. This allows what Kiayias and Lazos [17] call *accountability*: a blockchain governance system satisfies the property of accountability if whenever participants bring in a change, they are held individually responsible for it in a clearly defined way by the platform.

Hanson's futarchy proposal [15] pioneered the idea of using prediction markets to guide policy decisions, leveraging the collective intelligence of the market to evaluate which policy would produce better outcomes. Decision Scoring Rules [28, 3] and Decision Markets [29, 5] formalize it by conditioning monetary transfers on the observed outcome, but they are designed for single experts or community-wide market participation. In practice, decision markets have been deployed in DAOs either as the decision mechanism itself (MetaDAO [22]) or as a filtering layer that delegates the final decision to a committee (DAOstack [7]). An analysis on the latter noted that the system became self-referential: the collective intelligence of the market never showed up, and proposals were passed not because they were deemed

relevant, but because their authors wished to see them approved [9]. One could envision an inverted DAOstack architecture: the committee decides, and the outcome is evaluated ex-post by a market or a peer prediction method such as the one of Miller et al. [24]. This paper formalizes a new mechanism structure where transfers can be conditioned on both reports and impact of the outcomes, abstracting practical considerations such as the medium of exchange, the committee selection process, and the outcome evaluation procedure, allowing the incentive analysis to apply across diverse instantiations.

Mechanism Overview

We focus on binary decisions, the most common type in DAOs: approximately 67% of analyzed DAO proposals use a binary voting pattern [36].

Each committee member – hereafter *expert* – possesses two independent, private types of information: *idiosyncratic preferences*, how much they personally stand to gain or lose from each alternative (what they want), and *subjective beliefs*, their assessment of which alternative will improve the organization’s success metric (what they know). The designer wants to decide based on beliefs alone, but experts have no incentive to report their beliefs truthfully – their preferences create an incentive to misrepresent. Rather than trying to elicit (extract) preferences and beliefs separately, our mechanism induces each expert to submit a single continuous report that encodes both, and maximizes the probability that the aggregation of the reports matches the one the designer would make if all beliefs were observable.

Unlike Decision Scoring Rules, which elicit beliefs by directly scoring prediction accuracy, our mechanism ties payments to the realized outcome, so that each expert’s belief is revealed through their own expected-utility maximization. Incentive compatibility comes from the pivotal structure, an instance of the Groves Mechanism [13], in which each agent’s payment equals the externality their report imposes on the others. The role of the outcome-contingent rewards is to reshape the content of the true report, folding belief information into the scalar message that the pivotal mechanism would otherwise elicit as pure idiosyncratic preference. When experts do not have idiosyncratic preferences (a standard assumption in Decision Scoring Rules), the mechanism selects the alternative that a designer would select if given access to all private beliefs and aggregated them as a linear opinion pool, with each expert’s weight equal to $\frac{b}{1+p(a-1)}$, where p is the expert’s belief that their preferred alternative will succeed and a and b are positive numbers that can be different between experts. These parameters are the intuitive counterparts of the reward coefficients (α, β) derived in Section 4.

The mechanism sidesteps the strict-properness impossibilities of decision scoring rules under deterministic decision rules [29, 28] and ensures that each expert’s report is strictly optimal whenever they assign non-zero probability to any aggregate of the other agents’ reports.

Several mechanisms have been proposed for aggregating preferences in binary decision settings, each with different trade-offs regarding incentive compatibility, information elicitation, and budget requirements. However, most of these mechanisms do not account for the distinction between experts’ idiosyncratic preferences and their beliefs about which alternative benefits the organization. Table 1 compares decision-making mechanisms along the following dimensions: whether the mechanism elicits experts’ subjective beliefs about the success of each alternative (*Belief Elicitation*); whether it mitigates the influence of experts’ private preferences on the collective decision (*Reduce Idiosyncratic Noise*); whether it supports multiple simultaneous participants (*Multi-Agent*) and whether it operates in a single round (*One-Shot*). In this paper, we construct a mechanism comprising an allocation

rule, a pivotal transfer rule, and an outcome-contingent reward rule that depends on a boolean signal observed after the decision is implemented. We formulate a design objective for committee decision-making – maximizing the probability that the mechanism’s decision matches a *linear opinion pool* of the experts’ private beliefs, framing the designer’s problem as classification rather than welfare maximization. We design an affine parameterization of the reward rules and derive restrictions of the feasible region of design parameters that simultaneously satisfy dominant-strategy incentive compatibility, interim individual rationality, and budget feasibility. We derive the dominant strategy, in which each expert’s report jointly encodes preferences and beliefs into a single scalar, with the reward parameters controlling how strongly beliefs enter the report and how experts’ messages are scaled. Finally, we evaluate the mechanism via Monte Carlo simulations, showing that when experts receive signals correlated with the true success probabilities it outperforms majority voting with outcome-contingent rewards and a DSR applied to a random expert in most scenarios.

■ **Table 1** Comparison of decision-making mechanisms in decentralized organizations. ● = satisfied, ○ = partially satisfied, ○ = not satisfied. [†]Outcome-contingent subsidies can indirectly mitigate idiosyncratic noise, but these mechanisms were not designed for this purpose.

Feature	This	MV	QV	DSR	DM
Belief Elicitation	●	●	●	○	○
Reduce Idiosyncratic Noise	● [†]	● [†]	●	● [†]	● [†]
Multi-Agent	○	●	●	●	●
One-Shot	●	○	●	●	○

The remainder of the paper is organized as follows: Section 2 reviews related work. Section 3 presents the model, the mechanism structure, and the designer’s objective. Section 4 progressively derives the admissible reward rules that satisfy a set of important properties. Section 5 analyzes the design parameters and presents simulation results. Section 6 concludes.

2 Related Work

This paper sits at the intersection of three areas: governance in decentralized autonomous organizations, information aggregation in committees, and outcome-conditioned elicitation mechanisms (decision scoring rules and decision markets). DAO governance has identified the problem we address: committees making decisions where members hold both idiosyncratic preferences and private beliefs about what best serves the organization. We review information aggregation in committees and mechanisms with similar structure to ours.

Committees

A foundational result in committee decision-making is the Condorcet Jury Theorem (1785): if each member of a group independently identifies the correct alternative with probability greater than $1/2$, then majority voting selects the correct alternative with probability approaching 1 as the group grows [10]. The theorem assumes identical preferences and sincere voting. Austen-Smith and Banks [1] find that honest voting by all individuals is not rational even when they have common preferences. Li et al. [21] study committee decisions where members have conflicting preferences but share a common interest in making the correct choice, an objective close to ours. In their model, preferences are publicly known while private information (a signal about which alternative is correct) must be elicited through reports. The members’ preferences derive from the personal costs of incorrect decisions, and

the model does not distinguish between bias manifested in prior beliefs and bias manifested in preferences. They show that manipulation incentives persist under any decision rule – even with continuous reports – so no equilibrium fully aggregates members’ information, though some sharing still occurs. Nitzan and Paroush [27] studied the optimal weighted majority rule for a group of individuals with identical objectives but heterogeneous abilities. Each individual votes for one of two alternatives, identifying the correct one with a known probability that may differ across individuals.

Our mechanism departs from these settings by using transfers to establish a dominant strategy for each agent. The known equilibrium strategy allows the designer to provide formal incentive guarantees and to tune the mechanism’s parameters to control how private information enters the aggregate, even without observing individual types.

Decision Scoring Rules and Decision Markets

A DSR pairs a scoring function that rewards a single expert based on their report and the actual outcome with a decision rule that selects an action from the report; in decision markets, experts instead produce forecasts conditional on each action. Both were first introduced by [29]. Oosterheld and Conitzer [28] showed that DSRs can only strictly incentivize reporting the expected utility of the recommended action, not the full probability distribution over outcomes. Chen et al. [5] showed that in decision markets it is impossible to strictly incentivize truthful reports about the likelihood of achieving some outcomes under deterministic decision rules; in other words, to strictly incentivize truthful reports the decision maker must risk choosing a suboptimal alternative. The authors further study the case where a single expert simply recommends an action, extending the model introduced by [29].

Our setting departs from the single-expert focus of decision scoring rules: we aim to follow the action favored by the *aggregate* of the experts’ beliefs rather than that of any single expert. Extending DSRs to multiple agents directly introduces strategic manipulation. As in [29], we restrict attention to two possible outcomes, “good” and “bad”. Our mechanism assigns each expert a reward that differs according to whether the realized outcome was good or bad, and charges the pivotal transfer on their report. This makes truthful reporting a dominant strategy, strictly optimal whenever an expert assigns non-zero probability to every possible aggregate of the other experts’ reports. As a result, the designer can select the action it would have chosen with direct access to the experts’ private beliefs under a linear pooling rule.

Considering experts with their own preferences, [3] addresses self-interested experts by introducing compensation rules that offset the expert’s inherent utility from the decision. However, constructing proper compensation rules requires the principal to know the expert’s utility function; when this knowledge is unavailable, properness cannot be guaranteed in general. Our mechanism does not assume knowledge of experts’ preferences, though a known type distribution lets the designer anticipate its expected accuracy.

Two-Stage Mechanisms

The idea of conditioning transfers on post-decision information is not new. Mezzetti [23] proposed a two-stage Groves mechanism where agents report realized payoffs after the outcome, and [26] strengthened this to strict ex-post incentive compatibility. Our mechanism shares the two-stage structure but departs in two ways: the objective is classification rather than efficiency, and the second stage does not require agents to report their realized payoffs – a single boolean $\Delta \in \{-1, 1\}$ is observed, eliminating the need for incentive-compatible second-stage strategies.

3 Model

This section presents the model and the necessary background. We consider a single decision (one-shot setting) in which a DAO faces a choice between two mutually exclusive alternatives $\{A, B\}$. The committee consists of a set $N = \{1, \dots, n\}$ of risk-neutral experts who aim to maximize their expected utility. Following the implementation of the alternative, a boolean state of the world, $\Delta \in \{-1, 1\}$ is observed, where $\Delta = 1$ denotes a positive impact and $\Delta = -1$ denotes a negative impact. Only the implemented alternative is realized, so the mechanism observes Δ for the chosen option alone and never the counterfactual outcome of the discarded one; conditioning transfers on Δ requires only the realized impact of the chosen alternative to be resolvable to a boolean against a benchmark, not that the two alternatives be compared directly. We take the procedure that produces Δ – a community retrospective, a quantitative KPI threshold such as treasury growth or total value locked, or an arbitration method such as Kleros [20] – as given. Its choice is orthogonal to the incentive analysis, but the guarantees hold only if Δ is observed honestly (and experts understand how it is determined).

Experts have quasi-linear utility: a payment of x to the mechanism reduces the expert's utility by exactly x , with no indirect benefit from the collective pool increasing. The medium of exchange is left unspecified – it may be tokens from the DAO treasury, reputation points, or any other transferable unit of value. We assume that transfers are enforceable, e.g. via smart contracts, and that posting collateral does not change the expert's utility function. Experts are risk neutral, thus they value the payment now as much as they will value it later. The rules of the mechanism are common knowledge. We focus on environments where the number of experts is insufficient for traditional decentralized prediction markets to function effectively.

3.1 Information Structure

Each expert $i \in N$ has a private type – a collection of private information that fully characterizes the expert – comprising their idiosyncratic preferences and subjective beliefs.

The expert's idiosyncratic utility is captured by the pair $v_i = (v_i^A, v_i^B) \in \mathbb{R}^2$, where v_i^A and v_i^B represent the expert's personal utility derived from the implementation of option A and option B, respectively. Since exactly one of these options will be implemented, the expert can normalize their preference parameters accordingly so without loss of generality, at least one of these alternatives yields a non-negative idiosyncratic utility for the expert.

The expert's subjective beliefs are represented by the pair $p_i = (p_i^A, p_i^B) \in \mathcal{P}^2$, where $\mathcal{P} = (0, 1)$ and p_i^A and p_i^B denote the expert's private assessment of the probability that alternatives A and B will lead to a positive outcome for the organization, respectively. We exclude the cases where the experts are certain, to avoid unnecessary notation. They are formally defined as the conditional probabilities of observing a positive outcome given the chosen alternative: $p_i^A = \mathbb{P}_i(\Delta > 0 \mid A)$ and $p_i^B = \mathbb{P}_i(\Delta > 0 \mid B)$. These are subjective conditionals: they may differ across experts and generally will not coincide with the true success probabilities.

The expert's type is given by $\theta_i \in \Theta_i$, where $\theta_i = (v_i^A, v_i^B, p_i^A, p_i^B)$. We denote the type profile of all experts' types by $\theta = (\theta_1, \dots, \theta_n)$ and the belief profile by $\mathbf{p} = (p_1, \dots, p_n)$ with $p_i = (p_i^A, p_i^B)$. We operate under the Independent Private Values (IPV) model: each expert's type θ_i is drawn independently from some distribution over the type space Θ_i .

To clarify this assumption, we contrast our setting with the mineral rights auction [16]. There, each bidder receives a private geological signal, but the true value of the tract is a common function of all signals. Learning another bidder's signal would cause a bidder

to revise their own valuation. A correlated values model is necessary precisely because no public deliberation stage resolves the common-value uncertainty before bids are placed. In DAO governance the mechanism is preceded by forum discussions, temperature checks, and community calls that generate a common public signal observed by all experts. Each expert absorbs this shared information and updates accordingly. We capture what stays private after discussion. In this way, an expert's private signal carries no information about another expert's signal: learning it would not cause a further update.

We do not assume that the two pairs (v_i^A, v_i^B) and (p_i^A, p_i^B) are correlated. Experts tend to benefit from the options they believe will succeed which aligns private incentives with informational contributions and makes the designer's task easier. Independence is therefore a conservative assumption.

The mechanism takes the committee as given: the selection procedure determines robustness to coordination among experts. The properties rely on independent private values; when experts coordinate before reporting – through shared information, bribery, or vote-buying – this assumption breaks down.

3.2 Mechanism Structure

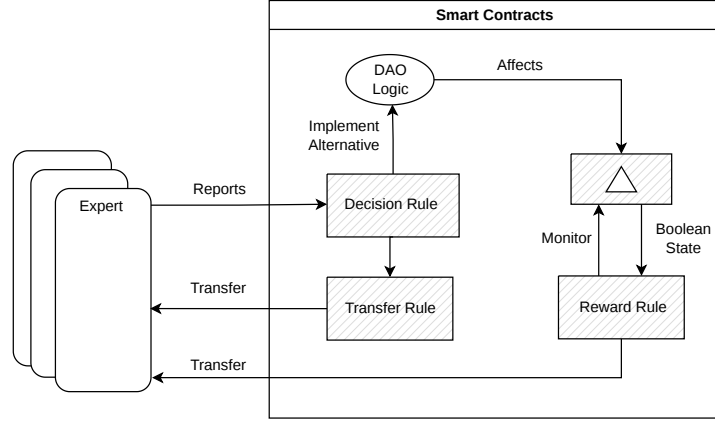
A direct revelation mechanism asks each participant to report their private information and then determines an outcome and monetary transfers as a function of the submitted messages. Formally, it consists of two components: an *allocation rule*, which maps reports to an outcome, and a *transfer rule*, which determines the payment each participant makes or receives based on the reports. We introduce a third component, a *reward rule*, that distributes payments after the outcome is observed. The reward rule is defined separately from the transfer rule because the expert evaluates it only in expectation over the unknown outcome at the time of reporting, while the transfer is known. This separation allows us to restrict the admissible reward rules independently of the allocation and transfer rules.

In our case, the mechanism proceeds in two stages, illustrated in Figure 1. First, each expert submits a message encoding how strongly the expert favors one alternative over the other. The mechanism aggregates these messages, selects an alternative, and charges each expert a transfer that depends only on the submitted messages. Second, after the outcome is observed, the mechanism distributes a reward that depends on whether the decision led to a positive or negative outcome. The expert's payoff is therefore determined by three factors: which option is chosen, how much they pay upfront, and what reward they receive once the outcome is revealed.

All experts share a common message space $\mathcal{R}$. The profile of messages submitted by all experts is denoted by $\mathbf{m} = (m_1, \dots, m_n) \in \mathcal{R}^n$. We use the notation $\mathbf{m}_{-i}$ for the profile of all experts except i , and sometimes write $\mathbf{m} = (m_i, \mathbf{m}_{-i})$.

We formally define a mechanism in this setting as a triplet $\mathcal{M} = (x, t, r)$. We use $\mathbf{t}, \mathbf{r}$ to refer to vectors denoting the application of each rule:

- $x : \mathcal{R}^n \rightarrow \{A, B\}$ is the deterministic allocation rule representing the collective decision.
- $\mathbf{t} = (t_1, \dots, t_n)$ is the transfer rule, where $t_i : \mathcal{R}^n \rightarrow \mathbb{R}$ represents a monetary transfer to expert i based solely on the submitted messages. A positive transfer means the expert receives money, while a negative transfer means they pay.
- $\mathbf{r} = (r_1, \dots, r_n)$ is the reward rule, where $r_i : \mathcal{R}^n \times \{-1, 1\} \rightarrow \mathbb{R}$ represents the ex-post reward distributed to expert i , contingent on the realized state of the world Δ .



■ **Figure 1** Mechanism Overview.

Given a type $\theta_i = (v_i^A, v_i^B, p_i^A, p_i^B)$ and a profile of submitted messages $\mathbf{m}$, let $o = x(\mathbf{m}) \in \{A, B\}$ denote the chosen alternative. We write v_i^o , p_i^o for the expert's utility and belief associated with the chosen alternative. The expected utility of expert i reflects the three factors described above:

$$\mathbb{E}[u_i(\mathbf{m} \mid \theta_i)] = v_i^o + t_i(\mathbf{m}) + \mathbb{E}_\Delta[r_i(\mathbf{m}, \Delta) \mid p_i^o] \quad (1)$$

3.3 Background: The Pivotal Mechanism

Our mechanism builds on the Pivotal Mechanism [13]. Each expert submits a scalar $m_i \in \mathbb{R}$ representing the intensity of their preference. The allocation rule selects A if the sum of messages is non-negative, otherwise it selects B :

$$x(\mathbf{m}) = \begin{cases} A & \text{if } \sum_{i=1}^n m_i \geq 0 \\ B & \text{otherwise} \end{cases} \quad (2)$$

The transfer charges each pivotal expert the externality they impose on others:

$$t_i(\mathbf{m}) = \begin{cases} -\sum_{j \neq i} m_j & \text{if } -m_i > \sum_{j \neq i} m_j > 0 \\ \sum_{j \neq i} m_j & \text{if } -m_i < \sum_{j \neq i} m_j < 0 \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

Non-pivotal experts pay nothing and reporting $m_i = v_i^A - v_i^B$ is a dominant strategy.

In the pivotal mechanism, each expert has a *winning set*: the set of profiles where their preferred alternative is selected. An expert who prefers A ($m_i > 0$) wins when $\sum_j m_j \geq 0$; submitting a larger positive message pushes the aggregate toward their winning set. An expert who prefers B ($m_i < 0$) wins when $\sum_j m_j < 0$; submitting a more negative message pushes toward theirs. The *critical value* is the minimum message needed to enter the winning set: $c_i = -\sum_{j \neq i} m_j$. The pivotal transfer (3) charges exactly this value to pivotal experts. By Myerson's lemma [25], a mechanism is DSIC if and only if the allocation is monotone in the message and each agent pays their critical value. DSIC is characterized in terms of the sum of both transfers: $t_i(\mathbf{m}) + \mathbb{E}_\Delta[r_i(\mathbf{m}, \Delta)]$.

3.4 Designer's Objective

Unlike most mechanisms that seek to maximize the sum of idiosyncratic preferences v_i , here these valuations are exactly the idiosyncratic noise the designer wants to filter out. The designer instead wants to select the alternative most likely to yield a positive outcome for the organization, which is given by the experts' beliefs, a quantity no agent wants to report directly.

The designer's goal is to select the alternative that an informed designer would choose if given direct access to all experts' private beliefs about each alternative's success probability. Let $x^* : \mathcal{P}^{2n} \rightarrow \{A, B\}$ denote an ideal decision rule – the allocation the designer would choose if all beliefs were observable. However, the designer does not observe p_i ; instead, the mechanism collects messages m_i that encode both v_i and p_i . The designer seeks to maximize the probability that the mechanism's decision agrees with the ideal one:

$$\max_{x, t, r} \mathbb{P}(x(\mathbf{m}^*) = x^*(\mathbf{p})), \quad (4)$$

where $\mathbf{m}^* = (m^*(v_1, p_1), \dots, m^*(v_n, p_n))$ is the message profile under the dominant strategy m^* induced by (x, t, r) , and the probability is over the random types (v_i, p_i) .

A *strategy* for agent i is a function $m : \mathbb{R}^2 \times \mathcal{P}^2 \rightarrow \mathbb{R}$ mapping the agent's type (v_i, p_i) to a message $m_i = m(v_i, p_i)$. The designer seeks a mechanism in which each agent has a dominant strategy that depends on both components of the type. Since the mechanism's instruments (x, t, r) cannot depend on p_i directly, the dominant strategy must be *implementable*: it must arise from the agent's optimization under rules independent of p_i .

We instantiate the general objective by choosing an ideal decision rule x^* based on *linear pooling* [37, 32]. Linear pooling is a well-studied method for combining expert forecasts: it averages the individual probability assessments with weights set by the designer. In our setting, this amounts to selecting A whenever the experts' average belief differential favors it:

$$x^*(\mathbf{p}) = \begin{cases} A & \text{if } \sum_{i=1}^n w_i^A p_i^A - w_i^B p_i^B > 0 \\ B & \text{otherwise} \end{cases} \quad (5)$$

An alternative is logarithmic pooling, which gives disproportionate weight to extreme assessments: if one expert assigns 90% to A while five others assign 40%, logarithmic pooling selects A , whereas linear pooling selects B .

Since the designer does not observe (p_i^A, p_i^B) , the mechanism must rely on the messages m_i . The designer therefore seeks a reward rule r such that the sign of the aggregate message matches the sign of the aggregate belief differential:

$$\text{sgn} \left(\sum_{i=1}^n m_i \right) = \text{sgn} \left(\sum_{i=1}^n w_i^A p_i^A - w_i^B p_i^B \right), \quad \forall \theta \in \Theta \quad (6)$$

subject to the constraints defined in the next section. We consider equal weights throughout the rest of the paper.

3.5 Formalization of the Properties

The desired properties for the decision-making mechanism are formalized below.

The mechanism must remain invariant to the relabeling of alternatives. Since the reports are reflected by the experts' utility functions we formalize this requirement as:

► **Property 1** (Symmetry). *A mechanism satisfies Symmetry if, for every agent i , every type $\theta_i = (v_i^A, v_i^B, p_i^A, p_i^B)$ with relabeled type $\tilde{\theta}_i = (v_i^B, v_i^A, p_i^B, p_i^A)$:*

$$\mathbb{E}[u_i(\mathbf{m} \mid \theta_i)] = \mathbb{E}[u_i(-\mathbf{m} \mid \tilde{\theta}_i)]$$

where $-\mathbf{m}$ is the profile obtained by negating each message.

Next, we establish the condition under which experts are willing to participate in the DAO's decision-making process, formally known as Individual Rationality (IR). We adopt the notion of interim IR: each expert knows their own preferences and beliefs when deciding whether to vote, but does not know what the other experts will report. Additionally, there also exist the notions of ex-ante (before learning one's type) and ex-post (after all messages are submitted).

► **Property 2** (Interim Individual Rationality). *A decision-making mechanism satisfies Interim Individual Rationality (IR) with reservation level $\bar{u}$ if:*

$$\forall i \in N, \forall \theta_i \in \Theta_i, \forall \mathbf{m}_{-i} \in \mathcal{R}_{-i} : \quad \mathbb{E}[u_i(m_i^*, \mathbf{m}_{-i} \mid \theta_i)] \geq \mathbb{E}[u_i(\mathbf{m}_{-i} \mid \theta_i)] + \bar{u}$$

for some $m_i^* \in \mathcal{R}$.

A committee decision is *non-excludable*: expert i is bound by the collective outcome whether or not they vote. The relevant outside option is therefore not zero utility but *abstention* – the expert declines to influence the vote while the others still decide, receiving $\mathbb{E}[u_i(\mathbf{m}_{-i} \mid \theta_i)]$, in which $t_i = 0$, the reward vanishes, and they obtain only their valuation $v_i^x(\mathbf{m}_{-i})$ of the outcome chosen by others. Participation is individually rational when the expected gain over this abstention payoff clears a reservation level $\bar{u} \leq 0$, which measures how far below the abstention payoff an expert will still tolerate in order to keep their seat. When exit is costless, an expert participates only if doing so is at least as good as abstaining, so $\bar{u} = 0$.

Having established the conditions for expert participation, we must also consider the financial viability of the mechanism from the perspective of the DAO. A mechanism satisfies the Budget Constraint if the total net compensation issued by the DAO to the experts is bounded by a maximum subsidy. This bound applies regardless of the medium of exchange: even when transfers are denominated in reputation points rather than tokens, the mechanism is constrained by the maximum value that experts attribute to reputation.

► **Property 3** (Budget Feasibility). *A decision-making mechanism satisfies Budget Feasibility with budget $c \geq 0$ if:*

$$\forall \mathbf{m} \in \mathcal{R}^n, \forall \Delta \in \{-1, 1\} : \quad \sum_{i \in N} (t_i(\mathbf{m}) + r_i(\mathbf{m}, \Delta)) \leq c$$

A mechanism is incentive compatible if experts maximize their expected utility by honestly revealing this information rather than behaving strategically. A DSIC mechanism guarantees that reporting honestly is an expert's best strategy regardless of the reports of other participants.

► **Property 4** (Dominant Strategy Incentive Compatibility). *A decision-making mechanism is Dominant Strategy Incentive Compatible (DSIC) if,*

$$\forall i \in N, \forall \theta_i \in \Theta_i, \forall m_i \in \mathcal{R}, \forall \mathbf{m}_{-i} \in \mathcal{R}^{n-1} :$$

$$\mathbb{E}[u_i(m_i^*, \mathbf{m}_{-i} \mid \theta_i)] \geq \mathbb{E}[u_i(m_i, \mathbf{m}_{-i} \mid \theta_i)]$$

where m_i^* denotes the honest reporting strategy for type θ_i .

Ideally, the designer would observe each expert's full type $\theta_i = (v_i, p_i)$, discard the preferences v_i , and decide based on the beliefs p_i alone. But experts have no incentive to report p_i truthfully – their preferences create an incentive to misrepresent their beliefs. Rather than eliciting v_i and p_i separately, we design a mechanism that induces experts to report a single number m_i^* that encodes both, in a way that brings the collective decision closer to the designer's goal. What counts as “truthful” depends on the question being asked. Throughout the paper, we use DSIC to mean that the mechanism has a dominant strategy $m_i^*(\theta_i)$. By the revelation principle, this is without loss of generality [12].

4 Mechanism Design Space

The mechanism uses the pivotal transfer from Section 3.3 and augments it with an outcome-contingent reward rule. The protocol proceeds as follows: (1) each expert deposits a stake into the governance smart contract; (2) each expert submits a message $m_i \in \mathbb{R}$; (3) the allocation rule selects an alternative; (4) the pivotal transfer is deducted from each expert's escrow; (5) after a predetermined evaluation period, the on-chain tool observes $\Delta \in \{-1, 1\}$; (6) the reward is distributed and remaining escrow is returned.

We restrict attention to *anonymous* reward rules, where every expert faces the same rule $r_i = r$. Non-anonymous rules, assigning per-expert parameters, let the designer attribute different weights to experts – prioritizing those with demonstrated expertise – without affecting the incentive guarantees. The properties from Section 3.5 progressively narrow the space of admissible reward rules.

Symmetry

The reward rule has four components: r_A^+ , r_A^- , r_B^+ , and r_B^- , corresponding to the two alternatives and two outcome realizations. These are expert i 's reward functions; throughout this paragraph we suppress the expert subscript. Since the mechanism must not favor either alternative, the reward for choosing A under a given profile must equal the reward for choosing B under the relabeled profile. This reduces the design space from four functions to two:

► **Proposition 5.** *The mechanism (x, t, r) satisfies Symmetry if and only if*

$$r_A^+(\mathbf{m}) = r_B^+(-\mathbf{m}), \quad r_A^-(\mathbf{m}) = r_B^-(-\mathbf{m}) \quad \forall \mathbf{m}$$

Proof. Let $o = x(\mathbf{m})$ denote the chosen alternative under a message profile $\mathbf{m}$ and $\bar{o} = x(-\mathbf{m})$. We have

$$\mathbb{E}[u_i(\mathbf{m} \mid \theta_i)] = v_i^o + t_i(\mathbf{m}) + p_i^o r_o^+(\mathbf{m}) + (1 - p_i^o) r_o^-(\mathbf{m})$$

$$\mathbb{E}[u_i(-\mathbf{m} \mid \tilde{\theta}_i)] = v_i^{\bar{o}} + t_i(-\mathbf{m}) + p_i^{\bar{o}} r_{\bar{o}}^+(-\mathbf{m}) + (1 - p_i^{\bar{o}}) r_{\bar{o}}^-(-\mathbf{m})$$

By the neutrality of the allocation rule (negating all messages selects the opposite alternative: $\bar{o} = x(-\mathbf{m})$ is the opposite of $o = x(\mathbf{m})$) and definition of the pivotal transfer, we have, $v_i^o + t_i(\mathbf{m}) = v_i^{\bar{o}} + t_i(-\mathbf{m})$. Since the relabeled type $\tilde{\theta}_i$ swaps $p_i^A \leftrightarrow p_i^B$, we have $p_i^{\bar{o}}$ under $\tilde{\theta}_i$ equals p_i^o under θ_i . Thus $\mathbb{E}[u_i(\mathbf{m} \mid \theta_i)] = \mathbb{E}[u_i(-\mathbf{m} \mid \tilde{\theta}_i)]$ iff $p_i^o r_o^+(\mathbf{m}) + (1 - p_i^o) r_o^-(\mathbf{m}) = p_i^{\bar{o}} r_{\bar{o}}^+(-\mathbf{m}) + (1 - p_i^{\bar{o}}) r_{\bar{o}}^-(-\mathbf{m})$ for all $p_i^o \in (0, 1)$. This holds iff $r_o^+(\mathbf{m}) = r_{\bar{o}}^+(-\mathbf{m})$ and $r_o^-(\mathbf{m}) = r_{\bar{o}}^-(-\mathbf{m})$, which gives $r_A^+(\mathbf{m}) = r_B^+(-\mathbf{m})$ and $r_A^-(\mathbf{m}) = r_B^-(-\mathbf{m})$ for all $\mathbf{m}$. ◀

17:12 Belief-Aware Pivotal Mechanism for DAO Committees

Symmetry thus collapses the four reward functions to two. We define $r_i^+(\mathbf{m}) := r_A^+(\mathbf{m})$ and $r_i^-(\mathbf{m}) := r_A^-(\mathbf{m})$, so that

$$r_i(\mathbf{m}, \Delta) = \begin{cases} r_i^+(\mathbf{m}) & \text{if } \Delta = 1 \\ r_i^-(\mathbf{m}) & \text{if } \Delta = -1 \end{cases}$$

It remains to determine how r_i^+ and r_i^- may depend on the message profile. We write them as functions of m_i and $\mathbf{m}_{-i}$; whenever the subscript i appears explicitly among the arguments, we drop it in the function name. The *expected transfer* of expert i – the transfer plus the reward evaluated in expectation over Δ according to the expert's belief p_i^o – is:

$$\hat{t}_i(\mathbf{m}) = t_i(\mathbf{m}) + p_i^o r^+(m_i, \mathbf{m}_{-i}) + (1 - p_i^o) r^-(m_i, \mathbf{m}_{-i}). \quad (7)$$

Budget Feasibility

The Budget Feasibility restricts how r^+ and r^- can grow with $|t_i|$:

► **Proposition 6.** *If the mechanism (x, t, r) satisfies the Budget Feasibility and the message space is unrestricted ($m_i \in \mathbb{R}$), then r^+ and r^- are bounded above by an affine function of $|t_i|$.*

Proof. For any fixed $\mathbf{m}_{-i}$, the transfer t_i can be made arbitrarily large in magnitude by varying m_i . If r^+ (or r^-) grows faster than $|t_i|$ in the positive direction, the total payout $\sum_{i \in N} (t_i + r_i^\pm)$ grows without bound, violating the Budget Feasibility. Hence the positive part of r^+ and r^- is at most affine in $|t_i|$. ◀

The proposition bounds the reward from above but not from below, and does so for each fixed $\mathbf{m}_{-i}$ separately. A priori, the admissible form is therefore:

$$\begin{aligned} r^+(m_i, \mathbf{m}_{-i}) &= (\alpha^+(\mathbf{m}_{-i}) - 1) t_i(m_i, \mathbf{m}_{-i}) + \beta^+(\mathbf{m}_{-i}) - g^+(m_i, \mathbf{m}_{-i}), \\ r^-(m_i, \mathbf{m}_{-i}) &= (\alpha^-(\mathbf{m}_{-i}) - 1) t_i(m_i, \mathbf{m}_{-i}) + \beta^-(\mathbf{m}_{-i}) - g^-(m_i, \mathbf{m}_{-i}) \end{aligned}$$

where $g^+, g^- \geq 0$ are arbitrary non-negative functions representing additional penalties beyond the affine component. Dominant-strategy incentive compatibility (Theorem 9) requires the induced transformation to be independent of $\mathbf{m}_{-i}$, which forces $\alpha^\pm$ and $\beta^\pm$ to be constant in $\mathbf{m}_{-i}$; otherwise the dominant strategy would depend on the other experts' reports. We set $g^+ = g^- = 0$, restricting to affine rewards, which implement the linear opinion pool that matches our instantiation of the designer's goal. Non-affine rewards remain DSIC whenever they correspond to a strictly increasing message transformation (Lemma 8). These functions might be used to implement a nonlinear aggregation rule; characterizing the non-affine family is left to future work. The resulting affine parameterization is:

$$r^+(m_i, \mathbf{m}_{-i}) = (\alpha^+ - 1) t_i(m_i, \mathbf{m}_{-i}) + \beta^+ \quad r^-(m_i, \mathbf{m}_{-i}) = (\alpha^- - 1) t_i(m_i, \mathbf{m}_{-i}) + \beta^- \quad (8)$$

where $\alpha^+, \alpha^-, \beta^+, \beta^- \in \mathbb{R}$ are design parameters. The coefficient $(\alpha^\pm - 1)$ measures how much the reward amplifies or attenuates the pivotal transfer when the outcome is positive or negative, and $\beta^\pm$ is a fixed payment distributed independently of the value transfer.

Substituting (8) into (7) and writing $\alpha(p) = \alpha^+ p + \alpha^- (1 - p)$ and $\beta(p) = \beta^+ p + \beta^- (1 - p)$, the expected transfer simplifies to:

$$\hat{t}_i(\mathbf{m}) = \alpha(p_i^o) \cdot t_i(\mathbf{m}) + \beta(p_i^o), \quad (9)$$

where $o = x(\mathbf{m})$ is the chosen alternative.

The β term acts as a belief-dependent shift of the expert's valuation; write $\hat{v}_i := v_i + \beta(p_i^A) - \beta(p_i^B)$ for the resulting *effective valuation*, where $v_i = v_i^A - v_i^B$. The expert behaves as a pivotal agent with valuation $\hat{v}_i$ in place of v_i .

The four parameters $(\alpha^+, \alpha^-, \beta^+, \beta^-)$ of the mechanism class (8) constitute the design space. The budget feasibility further restricts this space:

► **Proposition 7.** *Under the Budget Feasibility, the design parameters satisfy $\alpha^+ \geq 0$, $\alpha^- \geq 0$, $\beta^+ \leq c/n$, and $\beta^- \leq c/n$.*

Proof. Substituting (8) into the Budget Feasibility for $\Delta = 1$:

$$\sum_{i \in N} \left(t_i(\mathbf{m}) + (\alpha^+ - 1) t_i(\mathbf{m}) + \beta^+ \right) = \alpha^+ \sum_{i \in N} t_i(\mathbf{m}) + n\beta^+ \leq c \quad \forall \mathbf{m}.$$

If the expert is pivotal, $t_i(\mathbf{m}) \leq 0$ for all i , and $\sum_i t_i(\mathbf{m})$ can be made arbitrarily negative by increasing the magnitude of the messages. If $\alpha^+ < 0$, the term $\alpha^+ \sum_i t_i$ is positive and unbounded, violating the constraint. Hence $\alpha^+ \geq 0$. Given $\alpha^+ \geq 0$, the left-hand side is maximized when $\sum_i t_i = 0$ (no expert is pivotal), yielding $n\beta^+ \leq c$, i.e., $\beta^+ \leq c/n$. The same argument applied to $\Delta = -1$ gives $\alpha^- \geq 0$ and $\beta^- \leq c/n$. ◀

It remains to determine which parameter values $(\alpha^+, \alpha^-, \beta^+, \beta^-)$ of the mechanisms defined by (8) yield DSIC and IR, and what dominant strategy they induce.

Dominant Strategy Incentive Compatibility

The following lemma shows that applying a strictly increasing transformation to the message space of a DSIC mechanism preserves incentive compatibility, with the critical value transformed by the inverse.

► **Lemma 8.** *Let (x, t) be a DSIC mechanism with winning set W_i and critical value $c_i = \inf\{m_i : x(m_i, \mathbf{m}_{-i}) \in W_i\}$. If the designer wants to coerce an agent to submit $\varphi(v_i)$ instead of v_i , where $\varphi : \mathcal{R} \rightarrow \mathcal{R}$ is strictly increasing, then the transformed mechanism is DSIC with critical value $\varphi^{-1}(c_i)$.*

Proof. The original mechanism is DSIC, so the allocation rule $x(v_i, \mathbf{m}_{-i})$ is monotonically increasing in v_i and agents pay their critical value when in the winning set. Under the transformation, the allocation becomes $x(\varphi(v_i), \mathbf{m}_{-i})$, which is increasing in v_i since φ is strictly increasing. The new critical value is

$$c'_i = \inf\{v_i : \varphi(v_i) \in W_i(\mathbf{m}_{-i})\} = \inf\{v_i : \varphi(v_i) \geq c_i\} = \varphi^{-1}(c_i).$$

Since monotonicity is preserved and with critical value $\varphi^{-1}(c_i)$, the transformed mechanism is DSIC. ◀

In other words, if the designer applies a strictly increasing transformation φ to the critical value, then the mechanism remains DSIC and the agents have dominant strategy $\varphi^{-1}(v)$.

► **Theorem 9.** *Let $\alpha^+, \alpha^- \geq 0$ with $(\alpha^+, \alpha^-) \neq (0, 0)$. The mechanism (x, t, r) is DSIC with dominant strategy*

$$m^*(v_i, p_i) = \begin{cases} \frac{v_i + \beta(p_i^A) - \beta(p_i^B)}{\alpha(p_i^A)} & \text{if } v_i + \beta(p_i^A) - \beta(p_i^B) > 0 \\ \frac{v_i + \beta(p_i^A) - \beta(p_i^B)}{\alpha(p_i^B)} & \text{if } v_i + \beta(p_i^A) - \beta(p_i^B) < 0 \end{cases} \quad (10)$$

where $v_i = v_i^A - v_i^B$.

Proof. Fix expert i with beliefs $p_i = (p_i^A, p_i^B)$. The β term shifts the expert's effective valuation to $\hat{v}_i$; the α term scales the transfer by $\alpha(p_i^o)$. When $\hat{v}_i > 0$ the expert prefers A , so being in the winning set means $o = A$ and the scaling is $\alpha(p_i^A)$; when $\hat{v}_i < 0$ the expert prefers B , giving $\alpha(p_i^B)$. Define the transformation

$$\varphi(\hat{v}_i) = \begin{cases} \hat{v}_i / \alpha(p_i^A) & \text{if } \hat{v}_i > 0 \\ \hat{v}_i / \alpha(p_i^B) & \text{if } \hat{v}_i < 0 \end{cases}$$

which is strictly increasing when $\alpha(p_i^A), \alpha(p_i^B) > 0$. By Lemma 8, the mechanism is DSIC with dominant strategy $m_i^* = \varphi(\hat{v}_i)$, yielding (10). ◀

Note that φ must depend only on the expert's own type, not on $\mathbf{m}_{-i}$: if it did, the dominant strategy would depend on $\mathbf{m}_{-i}$ as well, and the mechanism would no longer be DSIC. The dominant strategy is weakly optimal for all type profiles and strictly optimal whenever the expert assigns non-zero probability to being pivotal, i.e., to the event that their report changes the collective decision.

When experts have no idiosyncratic preferences, the mechanism reduces to pure belief elicitation, let, $\Delta\beta := \beta^+ - \beta^-$,

► **Corollary 10.** *When $v_i = 0$ for all i and $\Delta\beta > 0$, the dominant strategy reduces to $m_i^* = \Delta\beta (p_i^A - p_i^B) / \alpha(p_i^o)$. In particular, taking $\alpha(p_i^o) = 1$ (e.g. $\alpha^+ = \alpha^- = 1$) recovers the equal-weight linear opinion pool of (5) exactly; other admissible α yield a belief-weighted pool with weights $1/\alpha(p_i^o)$.*

If the designer assigns non-anonymous reward rules, each expert i receives a per-expert $\Delta\beta_i$, which lets the designer weight experts by criteria such as expertise. Taking $\alpha = 1$, $m_i^* = \Delta\beta_i (p_i^A - p_i^B)$, the ratio $\Delta\beta_i / \Delta\beta_j$ fixes how much more expert i 's belief differential counts than expert j 's in the aggregate. For instance, setting $\Delta\beta_i = 2 \Delta\beta_j$ makes expert i pivotal-equivalent to two experts j holding the same belief.

Interim Individual Rationality

Using the reservation level $\bar{u}$ from Property 2.

► **Proposition 11.** *The mechanism satisfies Interim IR if and only if $\beta(p) \geq \bar{u}$ for all $p \in (0, 1)$.*

Proof. Under DSIC, truthful reporting weakly dominates any alternative report, including any report that makes expert non-pivotal. It therefore suffices to verify that the non-pivotal deviation satisfies IR; the truthful (possibly pivotal) report can only improve on it. When non-pivotal, $t_i = 0$ and the allocation is unchanged. The participation gain over $v_i^{x(\mathbf{m}_{-i})}$ is $\beta(p_i^{x(\mathbf{m})})$. IR requires this to be at least $\bar{u}$ for all $p \in (0, 1)$. Since $\beta(p) = \beta^+ p + \beta^- (1 - p)$ is linear in p , the condition $\beta(p) \geq \bar{u}$ for all $p \in (0, 1)$ holds if and only if $\beta^+, \beta^- \geq \bar{u}$. ◀

The four design parameters $(\alpha^+, \alpha^-, \beta^+, \beta^-)$ are constrained by the mechanism's structural and incentive properties. The resulting feasible region restrictions, are summarized in Table 2. The mechanisms of the form (8) leave the designer free to choose the following parameters $(\alpha^+, \alpha^-, \beta^+, \beta^-)$, where the ratio α^+ / α^- controls the reweighting of experts valuation and the β parameters determine the strength of the belief-based incentive.

■ **Table 2** Parameter restrictions from each mechanism property.

Property	Constraint	Reference
Budget Feasibility	$\alpha^+, \alpha^- \geq 0$ $\beta^+, \beta^- \leq c/n$	Proposition 7
DSIC	$(\alpha^+, \alpha^-) \neq (0, 0)$	Theorem 9
Interim IR	$\beta^+, \beta^- \geq \bar{u}$	Proposition 11
Feasible region restrictions	$\alpha^+, \alpha^- \geq 0, (\alpha^+, \alpha^-) \neq (0, 0)$ $\beta^+, \beta^- \in [\bar{u}, c/n]$	

A note on Complexity

The mechanism is computationally lightweight. The allocation rule computes a single sum $\sum_{i=1}^n m_i$ in $O(n)$ arithmetic operations. The Pivotal transfer for each expert requires evaluating $\sum_{j \neq i} m_j = \sum_{i=1}^n m_i - m_i$, which is $O(1)$ per expert given $\sum_{i=1}^n m_i$, yielding $O(n)$ total. The reward rule performs a constant number of operations per expert, thus the entire mechanism therefore runs in $O(n)$ time. The remaining cost is the evaluation of Δ , a single evaluation per decision. From a smart contract perspective, the on-chain cost is equally modest. The contract stores n messages and the budget c , computes the sum, determines the allocation, and records the Pivotal transfers – all in a single transaction with $O(n)$ storage writes and arithmetic. The ex-post reward is computed in a second transaction after Δ is resolved, requiring $O(n)$ operations. The only overhead beyond a standard voting mechanism is the evaluation of Δ , which depends on the choice of the evaluation tool, not on the mechanism itself.

5 Parameter Analysis

This section studies how the choice of parameters affects the mechanism’s decisions, how the β terms control the strength of the belief signal and how the ratio α^+/α^- reweights experts according to their beliefs. We further reformulate the designer’s choice as an optimization problem over the mechanisms of the form (8). The resulting objective has no closed form, so we evaluate it through Monte Carlo simulation, comparing the mechanism against majority voting with and without outcome-contingent rewards and a DSR applied to a random expert across committee sizes, preference biases, and belief-noise regimes.

5.1 Parameters

Each parameter pair shapes the dominant strategy in a distinct way: the β terms control the strength of the belief signal, while the α reweights experts according to their beliefs.

Betas

The belief differential $\beta(p_i^A) - \beta(p_i^B) = \Delta\beta(p_i^A - p_i^B)$, determines how strongly the expert’s message reflects their beliefs. When $\Delta\beta > 0$, experts who believe A is more likely to succeed ($p_i^A > p_i^B$) shift their message toward A , and symmetrically, experts who believe B is more likely to succeed ($p_i^B > p_i^A$) shift their message toward B . The magnitude $\Delta\beta$ controls the strength of this incentive in both directions.

Each expert's preferred-outcome belief p_i^o depends on the sign of $\hat{v}_i = v_i + \Delta\beta(p_i^A - p_i^B)$, so varying $\Delta\beta$ can flip which alternative an expert prefers, thereby changing the α weight. For this reason, widening $\Delta\beta$ can induce additional preference flips that alter the α -weights and may not always improve classification. Maximizing the belief signal is therefore not unconditionally optimal. In the environments we simulate, however, this effect does not bind: accuracy is monotone in $\Delta\beta$, so the optimum is attained at the budget bound $\Delta\beta = c/n - \bar{u}$ (i.e. $\beta^+ = c/n, \beta^- = \bar{u}$).

Alphas

Each expert's message is scaled by $1/\alpha(p_i^o)$, where p_i^o is the expert's belief that their preferred alternative will lead to a positive outcome. Since $\alpha(p)$ is linear in p , the relative values of α^+ and α^- determine how the mechanism reweights experts based on p_i^o . When $\alpha^+ < \alpha^-$, experts with high p_i^o have smaller $\alpha(p_i^o)$ and thus larger messages, contributing more to the aggregate. When $\alpha^+ > \alpha^-$, the opposite holds. When $\alpha^+ = \alpha^-$, all experts are weighted equally regardless of p_i^o .

► **Proposition 12.** *For any $\lambda > 0$, the mechanisms with parameters (α^+, α^-) and $(\lambda\alpha^+, \lambda\alpha^-)$ select the same alternative for every type profile.*

Proof. Under the dominant strategy, $m_i^* = \hat{v}_i/\alpha(p_i^o)$. Replacing (α^+, α^-) by $(\lambda\alpha^+, \lambda\alpha^-)$:

$$m_i^*|_{(\lambda\alpha^+, \lambda\alpha^-)} = \frac{\hat{v}_i}{\lambda\alpha^+p_i^o + \lambda\alpha^-(1-p_i^o)} = \frac{1}{\lambda\alpha^-} \cdot \frac{\hat{v}_i}{1 + p_i^o(\frac{\alpha^+}{\alpha^-} - 1)}$$

Since $\lambda, \alpha^- > 0$, the sign of $\sum_i m_i^*$ is preserved. Hence the decision depends on (α^+, α^-) only through the ratio α^+/α^- . ◀

When $\alpha^- > 0$, the decision can be parameterized by the ratio $\frac{\alpha^+}{\alpha^-}$. When $\frac{\alpha^+}{\alpha^-} < 1$, the denominator decreases with p , so experts with higher p_i^o contribute more to the aggregate. At $\alpha^- = 0$, the transfer is refunded when the outcome is negative, so the mechanism remains DSIC yet experts are never held accountable for pivotal decisions that lead to bad outcomes. Under the dominant strategy, the expert's net financial exposure is $\alpha^-|t_i|$ and the required escrow is invariant to α^- . When reducing α^- , experts compensate by submitting proportionally larger messages. The net exposure is always capped by the message the experts submit.

Parameterized Designer's Objective

The Designer's Problem from Section 3.4 – matching the sign of the aggregate message to the sign of the aggregate belief differential, as in (6) – can now be instantiated with the derived dominant strategy. The designer chooses the mechanism parameters $(\alpha^+, \alpha^-, \beta^+, \beta^-)$ to maximize the probability of correct classification for a given environment. The designer's problem:

$$\max_{\alpha^+, \alpha^-, \beta^+, \beta^-} \mathbb{P}_{v_i, p_i^A, p_i^B} \left(\text{sgn} \left(\sum_{i=1}^n \frac{v_i + \Delta\beta(p_i^A - p_i^B)}{\alpha(p_i^o)} \right) = \text{sgn} \left(\sum_{i=1}^n (p_i^A - p_i^B) \right) \right) \quad (11)$$

subject to

$$\begin{aligned} \alpha^+, \alpha^- &\geq 0, \quad (\alpha^+, \alpha^-) \neq (0, 0), \\ \beta^+, \beta^- &\in [\bar{u}, c/n] \end{aligned}$$

for some random distributions of the agents types.

5.2 Simulations

Since the belief p_i^o enters the denominator of each expert's dominant strategy, and the value of o depends on the sign of $\hat{v}_i$, the message aggregate is a sum of ratios of dependent random variables with no known closed form. We therefore evaluate the mechanism's performance via Monte Carlo simulation: for each parameter configuration, we repeatedly draw random type profiles from specified distributions, compute every expert's dominant-strategy message, apply the allocation rule, and report *accuracy*: the fraction of the drawn profiles on which the mechanism's decision coincides with the objectively superior alternative – the one with the higher true success probability.

A direct empirical comparison with decision markets is not well-posed in our setting. A decision market's realized accuracy is not a function of the type distribution alone; it depends on the market microstructure (trader's budget, risk attitude, the market scoring rule, and the number and order of trading rounds), none of which our mechanism requires. Moreover, under a deterministic decision rule no decision market admits a dominant strategy in the multi-agent setting; restoring strict properness requires a randomized decision rule [5]. Quadratic voting and the standard pivotal mechanism are also excluded, because neither is able to extract beliefs from experts.

We therefore compare our mechanism against two decision rules: *majority voting with rewards* (MV w/ rewards), where each expert votes for A if $\hat{v}_i > 0$ and the majority wins; and *majority voting without rewards* (MV), where each expert votes for A if $v_i > 0$. All three rules are evaluated on the same drawn type profiles.

We further compare against a decision scoring rule (DSR). A DSR pays a single expert a positive affine function of the principal's realized utility [28]; with a binary outcome and principal utility $u = 1$ for a good outcome and $u = -1$ for a bad one, this rewards the expert for the realized impact of the recommended action. Since DSR requires a single expert, we implement the DSR as a *random dictator*: one expert is drawn uniformly at random and their recommended alternative is taken.

Setup

We specify a parametric model for the experts' types, defining distributions for idiosyncratic preferences and belief noise. The true success probabilities are symmetric around 0.5, separated by a *quality gap* parameter $\delta \in \{0.05, 0.15, 0.25\}$, corresponding to hard, moderate, and easy classification tasks, capturing the case the designer typically cares about most: one alternative is likely to succeed while the other is likely to fail. When both alternatives would succeed (p^A, p^B both high) or both would fail (p^A, p^B both low), the stakes of choosing correctly are lower. By the Symmetry property, we can without loss of generality let A be the superior alternative.

Idiosyncratic preferences are drawn independently with mean $\mu_v \in \{0, -1\}$, where $\mu_v = 0$ is the unbiased case and $\mu_v = -1$ represents a committee that on average opposes the superior alternative. The variance is normalized to one, so $\Delta\beta$ is measured in units of preference intensity. Each expert receives a noisy private signal centered at the truth with standard deviation $\sigma \in \{0.2, 0.8, 2.0\}$ (*belief noise*), truncated to $(0, 1)$. A small σ corresponds to a well-informed committee, while a large σ represents dispersed information. Each expert's type is drawn as:

$$v_i \sim \mathcal{N}(\mu_v, 1), \quad p_i^A \sim \text{TN}(\tfrac{1}{2} + \delta, \sigma^2, 0, 1), \quad p_i^B \sim \text{TN}(\tfrac{1}{2} - \delta, \sigma^2, 0, 1),$$

where $\text{TN}(\mu, \sigma^2, a, b)$ denotes a truncated normal distribution on $(0, 1)$. Results with uniform preferences $v_i \sim U(\mu_v - 1, \mu_v + 1)$ are similar and omitted for brevity.

A grid search over $\alpha^+/\alpha^- \in [0, 2]$ reveals that the optimal ratio is consistently close to zero: the accuracy gap between $\alpha^+ = 0$ and the optimal α^+/α^- never exceeds 3% across all tested parameter configurations, so we adopt $\alpha^+ = 0$ as a default.

For each parameter configuration, we draw 10^6 independent type profiles. We vary $\Delta\beta \in [0, 5]$ on a grid of 30 equally spaced values and committee sizes $n \in \{5, 49\}$ (chosen odd to avoid ties in majority voting).

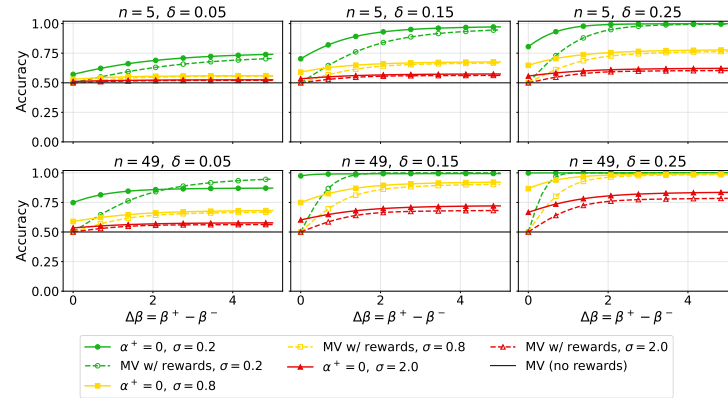
All simulations were implemented in Python with a fixed random seed. Table 3 summarizes the full set of varied parameters and their ranges.

■ **Table 3** Parameters varied in the Monte Carlo simulations.

Parameter	Symbol	Values
Quality gap	δ	$\{0.05, 0.15, 0.25\}$
Preference bias	μ_v	$\{0, -1\}$
Belief noise	σ	$\{0.2, 0.8, 2.0\}$
Committee size	n	$\{5, 49\}$
Belief budget	$\Delta\beta$	$[0, 5]$

Majority Voting

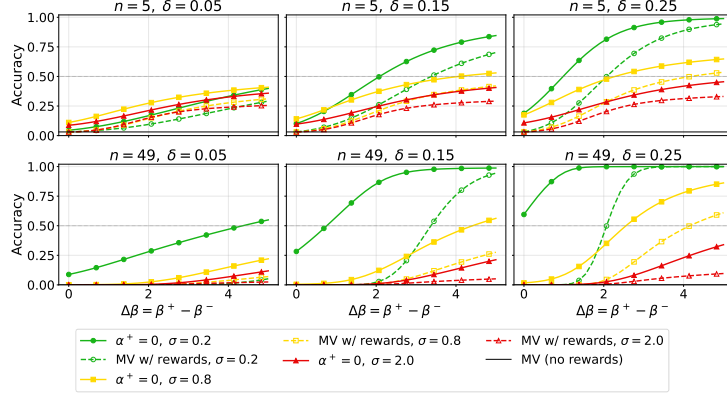
Figures 2 and 3 show that our mechanism outperforms majority voting across almost every parameter configuration, the exception being the small quality gap ($\delta = 0.05$) in the large committee ($n = 49$), where the two are comparable. The margin is widest when the belief budget $\Delta\beta$ is small.



■ **Figure 2** Unbiased committee ($\mu_v = 0$): classification accuracy as a function of $\Delta\beta$. Solid: our mechanism ($\alpha^+ = 0$). Dashed: MV with rewards. MV without rewards is the horizontal line at 0.5. Rows: committee size n . Columns: quality gap δ . Colors: belief noise σ . Preferences: $v_i \sim \mathcal{N}(\mu_v, 1)$.

At $\Delta\beta = 0$ the mechanism already reweighs experts through $1/\alpha(p_i^o)$, so beliefs scale each report even with no belief budget. For $\Delta\beta > 0$, messages are additionally pushed toward the expert's belief differential and accuracy rises steeply before flattening, as the belief signal comes to dominate preference noise. The quality gap δ and belief noise σ act as expected: a larger δ separates the true success probabilities and eases the task for every method, while a larger σ disperses the signals and degrades accuracy.

From a design standpoint, $\Delta\beta$ is bounded above by $c/n - \bar{u}$, so the designer can pick the minimum $\Delta\beta$ – and hence the minimum per-expert budget the DAO must commit – needed to reach a target accuracy.



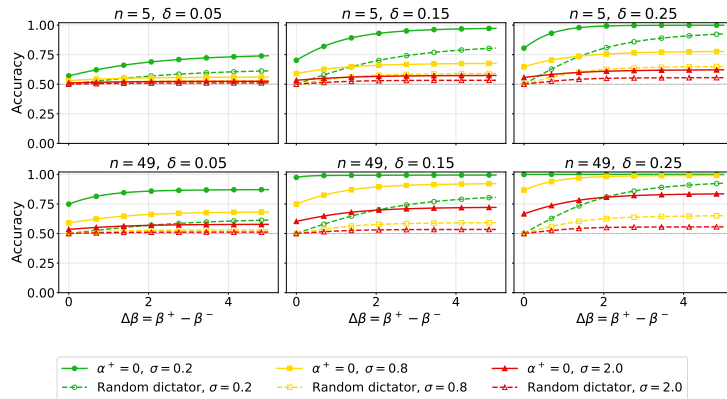
■ **Figure 3** Biased committee ($\mu_v = -1$), where members on average prefer the inferior alternative. Conventions as in Figure 2, except for the MV with no rewards that has a value close to 0.

For the biased committee (Figure 3), majority voting without rewards has approximately 0 accuracy, and with rewards it keeps only the sign of the shifted valuation, discarding the magnitude. Our mechanism instead scales each report by $1/\alpha(p_i^o)$, so an expert backing an option they believe is unlikely to succeed carries little weight in the aggregate.

Results with uniform preferences $v_i \sim U(\mu_v - 1, \mu_v + 1)$ are qualitatively similar, with one exception. At $\mu_v = -1$ the support is $[-2, 0]$, so every expert prefers B : at $\Delta\beta = 0$ no expert ever votes for A , reweighting alone cannot shift the decision, and accuracy is zero – unlike the normal case, whose tails leave some experts with $v_i > 0$. Once $\Delta\beta > 0$ the belief signal flips some experts toward A and the mechanism recovers.

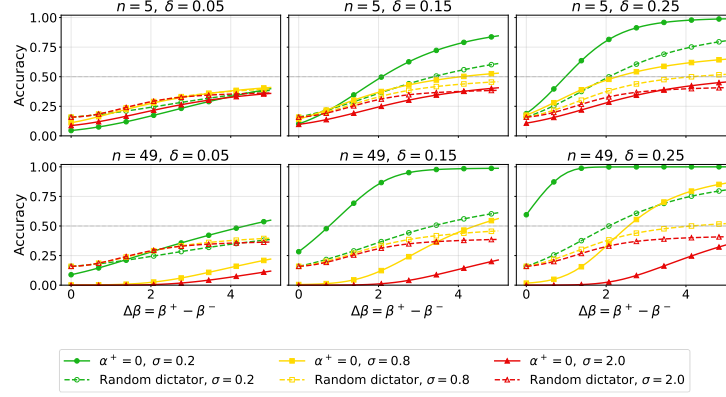
Random dictator

Figure 4 shows that in the unbiased committee our mechanism outperforms the DSR across every configuration.



■ **Figure 4** Unbiased committee ($\mu_v = 0$): classification accuracy of our mechanism ($\alpha^+ = 0$, solid) versus the random dictator with the same outcome-contingent incentives (dashed), as a function of $\Delta\beta$. Conventions as in Figure 2 σ . Preferences: $v_i \sim \mathcal{N}(\mu_v, 1)$.

In the biased committee (Figure 5) the DSR is more accurate at small $\Delta\beta$ in some scenarios.



■ **Figure 5** Biased committee ($\mu_v = -1$): our mechanism ($\alpha^+ = 0$, solid) versus the random dictator (dashed). Conventions as in Figure 2.

6 Conclusions

This paper presents a mechanism for committee decision-making with outcome-contingent transfers that satisfies dominant-strategy incentive compatibility, interim individual rationality, and budget feasibility. Each expert submits a single scalar that jointly encodes preferences and beliefs, and the mechanism maximizes the probability that its decision agrees with a linear opinion pool of the experts' private beliefs. The framework abstracts away the medium of exchange, the method of outcome evaluation, and the committee election process, providing a reference construction that practical designs can instantiate. The mechanism provides value even without a budget: the α reweighs experts, so informative beliefs can enter the aggregate, yielding higher classification accuracy than a standard pivotal mechanism at no additional cost. The budget amplifies this effect by adding an explicit belief shift to each expert's message.

When experts have no idiosyncratic preferences, the dominant strategy reduces to a scaled belief differential and the mechanism selects the alternative according to the linear opinion pool. This elicitation is strictly incentive compatible whenever each expert assigns non-zero probability to every possible aggregate of the other experts' reports. Decision scoring rules cannot achieve this in the multi-agent setting: under a deterministic decision rule, each expert's expected payoff depends on the collective decision, which in turn depends on the other experts' reports, so agents have an incentive to misreport to change the collective decision in their favor [29]. Restoring properness requires randomized decision rules [5], which introduce practical difficulties. Our mechanism sidesteps this impossibility by inducing a dominant strategy under a deterministic allocation rule and across multiple agents.

We focus on binary choices, which cover the majority of DAO proposals [36]. The designer's objective reduces to matching the sign of a linear opinion pool. Although the linear opinion pool generalizes naturally to k alternatives, the mechanism design side is less straightforward. With two alternatives, the classification boundary is a single threshold that maps to the pivotal mechanism's allocation rule. With $k \geq 3$, the designer must match an arg max over k pooled probabilities, and the interaction between outcome-contingent rewards and a multi-dimensional message space remains an open problem.

References

- 1 David Austen-Smith and Jeffrey S. Banks. Information aggregation, rationality, and the Condorcet jury theorem. *American Political Science Review*, 90(1):34–45, 1996.
- 2 Martin Becze, Hudson Jameson, et al. EIP-1: EIP purpose and guidelines. *Ethereum Improvement Proposals*, 27, 2015.
- 3 Craig Boutilier. Eliciting forecasts from self-interested experts: scoring rules for decision makers. *arXiv preprint arXiv:1106.2489*, 2011. [arXiv:1106.2489](https://arxiv.org/abs/1106.2489).
- 4 Vitalik Buterin, Zoë Hitzig, and E. Glen Weyl. A flexible design for funding public goods. *Manage. Sci.*, 65(11):5171–5187, November 2019. [doi:10.1287/mnsc.2019.3337](https://doi.org/10.1287/mnsc.2019.3337).
- 5 Yiling Chen, Ian Kash, Mike Ruberry, and Victor Shnayder. Decision markets with good incentives. In *International Workshop on Internet and Network Economics*, pages 72–83. Springer, 2011. [doi:10.1007/978-3-642-25510-6_7](https://doi.org/10.1007/978-3-642-25510-6_7).
- 6 Lin William Cong, Daniel Rabeti, Charles CY Wang, and Yu Yan. Centralized governance in decentralized organizations. *Available at SSRN 5168660*, 2025.
- 7 DAOstack. DAOstack whitepaper: An operating system for collective intelligence. <https://www.allcryptowhitepapers.com/daostack-whitepaper/>, 2018. Accessed: 2026-03-14.
- 8 Luke Dashjr. BIP 2: BIP process, revised. <https://github.com/bitcoin/bips/blob/master/bip-0002.mediawiki>, 2016. Bitcoin Improvement Proposal.
- 9 David Davó, Javier Arroyo, Samer Hassan, and Silvia Semenzin. The rise and fall of DAOstack: lessons for decentralized autonomous organizations. *PeerJ Computer Science*, 11:e3320, 2025. [doi:10.7717/PEERJ-CS.3320](https://doi.org/10.7717/PEERJ-CS.3320).
- 10 Nicolas de Condorcet. *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. Imprimerie Royale, Paris, 1785.
- 11 Nicola Dimitri. Voting in DAOs. *Distributed ledger technologies: Research and practice*, 2(4):1–12, 2023. [doi:10.1145/3624574](https://doi.org/10.1145/3624574).
- 12 Allan Gibbard. Manipulation of voting schemes: A general result. *Econometrica*, 41(4):587–601, 1973.
- 13 Jerry Green and Jean-Jacques Laffont. *Incentives in public decision-making*. Elsevier North-Holland, 1979.
- 14 Jiasun Han, Jongsub Lee, and Tao Li. DAO governance. *Journal of Corporate Finance*, 2025.
- 15 Robin Hanson. Shall we vote on values, but bet on beliefs? *Journal of Political Philosophy*, 21(2):151–178, 2013.
- 16 John H Kagel and Dan Levin. The winner's curse and public information in common value auctions. *The American economic review*, pages 894–920, 1986.
- 17 Aggelos Kiayias and Philip Lazos. SoK: blockchain governance. In *Proceedings of the 4th ACM Conference on Advances in Financial Technologies*, pages 61–73, 2022. [doi:10.1145/3558535.3559794](https://doi.org/10.1145/3558535.3559794).
- 18 Stefan Kitzler, Stefano Ballestti, Pietro Saggese, Hans-Martin Bernhard, and Markus Strohmaier. The governance of decentralized autonomous organizations: A study of contributors' influence, networks, and shifts in voting power. In *Financial Cryptography and Data Security (FC)*, 2024.
- 19 Steven P Lalley and E Glen Weyl. Quadratic voting: How mechanism design can radicalize democracy. In *AEA Papers and Proceedings*, volume 108, pages 33–37. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203, 2018.
- 20 Clément Lesaege, Federico Ast, and William George. Kleros: Short paper v1.0.7, 2019. URL: <https://kleros.io/whitepaper.pdf>.
- 21 Hao Li, Sherwin Rosen, and Wing Suen. Conflicts and common interests in committees. NBER Working Paper 7158, National Bureau of Economic Research, 1999.
- 22 MetaDAO. MetaDAO, n.d. Accessed: 2026-03-13. URL: <https://www.metadao.fi/>.
- 23 Claudio Mezzetti. Mechanism design with interdependent valuations: Efficiency. *Econometrica*, 72(5):1617–1626, 2004.

- 24 Nolan Miller, Paul Resnick, and Richard Zeckhauser. Eliciting informative feedback: The peer-prediction method. *Management Science*, 51(9):1359–1373, 2005. doi:10.1287/MNSC.1050.0379.
- 25 Roger B. Myerson. Optimal auction design. *Mathematics of Operations Research*, 6(1):58–73, 1981. doi:10.1287/MOOR.6.1.58.
- 26 Swaprava Nath and Onno Zoeter. A strict ex-post incentive compatible mechanism for interdependent valuations. *Economics Letters*, 121(2):321–325, 2013.
- 27 Shmuel Nitzan and Jacob Paroush. Optimal decision rules in uncertain dichotomous choice situations. *International Economic Review*, 23(2):289–297, 1982.
- 28 Caspar Oosterheld and Vincent Conitzer. Decision scoring rules. In *WINE*, page 468, 2020. URL: <https://link.springer.com/content/pdf/10.1007/978-3-030-64946-3.pdf#page=471>.
- 29 Abraham Othman and Tuomas Sandholm. Decision rules and decision markets. In *AAMAS*, pages 625–632, 2010. URL: <https://dl.acm.org/citation.cfm?id=1838288>.
- 30 Polkadot Wiki. Web3 and Polkadot, 2025. Accessed: 2026-03-12. URL: <https://wiki.polkadot.com/general/web3-and-polkadot/>.
- 31 Project Catalyst. Project catalyst documentation. <https://docs.projectcatalyst.io/>, n.d. Accessed: 2026-03-12.
- 32 Roopesh Ranjan and Tilmann Gneiting. Combining probability forecasts. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 72(1):71–91, 2010.
- 33 Olivier Rikken, Marijn Janssen, and Zenlin Kwee. Governance challenges of blockchain and decentralized autonomous organizations. *Information Polity*, 24(4):397–417, 2019. doi:10.3233/IP-190154.
- 34 T Sharma, Y Potter, K Pongmala, H Wang, A Miller, D Song, and Y Wang. Future of algorithmic organization: Large-scale analysis of decentralized autonomous organizations (DAOs). arXiv, 2024.
- 35 Nimrod Talmon. Social choice around decentralized autonomous organizations: On the computational social choice of digital communities. In *Proceedings of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 1768–1773, 2023. doi:10.5555/3545946.3598837.
- 36 Qin Wang, Guangsheng Yu, Yilin Sai, Caijun Sun, Lam Nguyen, and Shiping Chen. Understanding DAOs: An empirical study on governance dynamics. *IEEE Transactions on Computational Social Systems*, PP:1–19, January 2025. doi:10.1109/TCSS.2025.3539889.
- 37 Robert L. Winkler, Yael Grushka-Cockayne, Kenneth C. Lichtendahl, and Victor Richmond R. Jose. Probability forecasts and their combination: A research perspective. *Decision Analysis*, 16(4):239–260, 2019. doi:10.1287/DECA.2019.0391.