Delegated Control, Agency Conflicts and Endogenous Governance Cycles in Token-Based Platforms

Michele Fabi[*] Valerie Laturnus[†] Romain Rossello[‡]

January 16, 2026

EARLY VERSION - NOT FOR DISTRIBUTION

ABSTRACT

Token-based platform governance has evolved into a mediated voting system, in which tokenholders outsource voting to specialized agents. While this *vote delegation* can improve governance efficiency, it also reintroduces agency conflicts. We develop a dynamic principal-agent model in which agents are initially unknown and exert effort to build reputation. However, as perceived quality (reputation) rises, tokenholders rationally reduce active oversight. This decline in monitoring increases agents' returns to opportunism and misconduct, eventually leading to governance failures and agent turnover. To mitigate these governance cycles, we show that designs that slow reputation accumulation and use token-based compensation improve the platform's stability.

JEL classification: D72, D82, D86, G34, G23, L22

Keywords: Decentralized governance, Platform economics, Platform productivity, Liquidity constraint, Governance tokens, Voting rights, Vote delegation, Agency conflicts, Incentive compatibility, Contract design

[*]Telecom Paris-CREST (Institut Polytechnique de Paris), 5 Avenue Henry Le Chatelier, 91120 Palaiseau, France. ✉:Michele.FABI@ensae.fr.

[†]Durham University, Business School, Mill Hill Lane, Durham DH1 3LB, United Kingdom. ✉:valerie.laturnus@durham.ac.uk.

[‡]HEC Paris, 1 Rue de la Libération, 78350 Jouy-en-Josas, France. ✉:romain.rossello@hec.edu.

1 Introduction

Agency conflicts are universal in corporate governance, but organizations differ in how they mitigate them. For instance, some firms rely more on informal governance, such as trust, reputation, and peer discipline, while others use formal enforcement, including monitoring, board oversight, compliance, and the market for corporate control. What determines an organization's governance choice in mitigating agency costs, and how does it evolve over time? This paper studies these questions in a dynamic principal-agent framework.

From an economic perspective, one key mechanism is the information and enforceability environment. When actions are verifiable and contracts are enforceable, firms can rely on formal oversight instead of paying a premium for trustworthy managers. Conversely, when efforts are difficult to observe or specify in contracts, monitoring becomes costly and incomplete. In this case, governance places greater weight on screening, reputation, and other trust- and relationship-based mechanisms. Because both monitoring and trust-building are costly, most corporate governance systems operate between these extremes and substitute between formal and informal control as conditions change. The key economic problem is that agents respond to the incentives created by the governance environment. In other words, agents' reputations and trustworthiness are endogenous. They invest in reputational capital depending on the expected returns to opportunism.

To illustrate this mechanism, we investigate a setting where this trade-off is unusually transparent: the token-based platform governance (Abadi and Brunnermeier, 2024; Sockin and Xiong, 2023; Cong et al., 2022). Unlike traditional platforms, which are typically controlled by its founder, token-based platforms share ownership and control with its tokenholders (i.e., investors and platform users). These tokenholders can either incur the costs of becoming informed and monitoring directly, or they can delegate voting power to specialized, trustworthy intermediaries (delegates). Decision-making is frequent and highly technical, creating high fixed costs of attention and making delegation generally valuable. At the same time, many economically relevant actions are difficult to verify ex ante. Thus, even though actions and outcomes are recorded on blockchains, contractibility remains limited, and signaling, reputation and

trust play a central role in sustaining governance.

Our dynamic principal-agent model reveals a persistent, cyclical dynamic inherent to the governance environment. This cycle is driven by the interplay between the reputation of delegates (agents) and the monitoring incentives of tokenholders (principals). Initially, a newly selected delegate has limited reputational capital and therefore exerts high effort and behaves cooperatively to build credibility. As the perceived quality and trust in delegates increases, token holders rationally reduce monitoring. The resulting decline in oversight increases the returns to opportunism and the likelihood of self-dealing, governance failures, and delegate turnover. This resets reputations and restarts the cycle.

The model's findings lead to several practical implications for designing more robust decentralized governance systems. For instance, poorly designed compensation can amplify the cyclical dynamic by increasing the returns to short-term reputation farming. To align incentives, platforms should compensate delegates primarily with governance tokens that are locked and exposed to downside risk (e.g., vesting and slashing), so that delegates internalize the costs of governance failures. In parallel, governance processes should shift evaluation away from cheap, activity-based signals (proposal count, voting frequency) toward more decision-relevant disclosures (e.g., independent audits, and ex post performance scoring). This can slow the reputation-building of delegates who work hard at the beginning, only to artificially accelerate their status, which can lead to premature trust and the decline in monitoring.

Token-based platform governance, i.e, the use of cryptocurrencies or tokens for voting and collective decision-making, has emerged on blockchains that support the deployment of smart contracts (i.e, code executed on the blockchain). A growing literature highlights the condition-based enforceability of these contracts, which allows anonymous agents to commit credibly to financial transactions (see John et al. (2023) for a review). This enforceability contributed to the development of financially meaningful services (including exchange, saving, and lending) for blockchain participants, collectively referred to as decentralized finance (DeFi) and surveyed in (Harvey et al., 2021, 2024; Makarov and Schoar, 2022). More recently, it has also evolved into credible voting systems (Ferreira et al., 2023; Ferreira and Li, 2024).

The first major implementation of on-chain token voting was the investment collective known as "TheDAO", launched in April 2016. In this project, anonymous participants contributed funds and adopted a democratic, token-weighted voting mechanism (one token contributed per vote) to decide how the pooled capital would be allocated. The project collapsed shortly after launch because individual participants were able to drain funds by exploiting a coding vulnerability in the smart contract (see DuPont (2017) for a review). Despite its failure, the core concept of token-based platform governance has persisted. Beginning around 2019/2020, digital platforms that issued tokens to raise funds reintroduced and expanded the token-weighted voting to *vote delegations* (see Section 2 for institutional background information).

To provide a first impression, we study a sample of 31 token-based platform governance systems, and collect information on delegates, the distribution of voting rights among participants, voting behaviour and decision-making outcomes. In addition, we collect data on the size of platform treasuries and the frequency with which malicious agents intentionally misbehave. In this setting, those in power can implement platform policies designed to drain the funds (so-called platform's treasuries).

In general, we classify delegates into three broad categories, i.e., (i) professional organizations that supply governance services, (ii) intrinsically motivated individuals, and (iii) student groups and other nonprofit entities. Among our sample, we found more than 12 million active delegates participating, of whom only 100 are professional companies, 10 are student groups or other non-profit entities, and a potentially very large number are individuals. The majority of these addresses have no social media handle, so it is impossible to identify them. In large token-based governance systems, delegates actively compete for delegated voting power, and there is substantial turnover among pivotal delegates. In contrast, smaller governance systems have little delegation, and voting power remains in the hands of a few agents. Next, we examine delegation misconduct and find only 15 incidents in which delegates actively attempted to drain funds.

Next, we model the governance environment to study dynamic agency frictions between tokenholders and delegates in a decentralized platform. The analysis uses a continuous-time, infinite-horizon framework

with two classes of agents: tokenholders (principals) and delegates (agents). Delegates are heterogeneous. An honest delegate maximizes expected utility from compensation and token-value appreciation, whereas a malicious delegate has an additional option value, i.e., the ability to opportunistically attack the platform and drain the treasury.

Platform output (productivity) is observable and increases with effort. However, delegate effort itself is not directly observable or contractible. The platform's treasury evolves endogenously, i.e., it accumulates through platform fees and interest income and is depleted by delegate compensation or a successful attack. The governance token supply is stochastic and adjusts at discrete governance events to implement the compensation contract. Tokenholders face a costly information-acquisition problem. At governance events, they choose whether to incur an attention cost to become informed and identify the higher-quality delegate, or to vote based on noisy public signals. This setting generates an equilibrium trade-off between monitoring intensity and reliance on reputation.

A key implication of the model is the emergence of cyclical dynamics in platform governance. Initially, delegate reputation rises through Bayesian updating. As reputation increases, however, tokenholders' Marginal Benefit of Information (MBI) acquisition declines. Once reputation is sufficiently high, tokenholders optimally reduce or cease monitoring, creating a window in which a malicious delegate can attack the treasury. Counterintuitively, the model implies that malicious delegates may exert higher effort early on than honest delegates in order to inflate productivity and treasury value, thereby accumulating reputational capital and increasing the expected payoff from a later attack.

We also complement the analytical results with a series of simulations designed to quantify the dynamic implications of delegated governance under realistic parametrizations. Finally, the analysis suggests that these cycles can be mitigated by designing token-based compensation to raise the opportunity cost of attacks and by limiting artificial reputation inflation, so that monitoring incentives remain active throughout a delegate's tenure.

By endogenizing both reputation formation and monitoring incentives, the paper explains why governance structures differ across environments, and why the trade-off between trust and monitoring

evolves over time rather than remaining fixed. This is the main contribution of the paper.

A similar idea is investigated in related studies. Aghion et al. (2010) examines cross-country differences in regulation and shows that countries with more regulation tend to have lower social trust. The basic logic is that when citizens cannot easily monitor and discipline others, they demand more government regulation. But extensive regulation does not automatically build trust, so societies can end up in different regimes, i.e, high regulation with low trust versus low regulation with high trust.

We also study how the trade-off between trust (through reputation-building) and monitoring can lead to different governance regimes, however, in a different setting and with a different mechanism. In Aghion et al. (2010), the feedback runs from regulation to trust, i.e., with more regulation, people rely less on goodwill and mutual monitoring, so over time social trust can decline. In our model, the feedback runs from reputation to monitoring. As delegates build reputation, tokenholders increasingly trust them and rationally reduce monitoring. However, this trust can increase the returns to opportunism. As a result, governance failures in our setting arise from adverse selection and moral hazard under delegated control, rather than from broad societal preferences for regulation and trust.

Diamond (1984) studies delegated monitoring. When monitoring is costly, principals delegate it to an intermediary, but this creates a new agency problem, i.e., the intermediary must be given incentives to actually monitor. Similarly, we argue that, in our setting, delegates need the right incentives to commit to governance throughout their tenure, which makes opportunism unattractive. Furthermore, we introduce a formal framework that characterizes the governance subgame in token-based platforms with delegated voting, and we derive conditions under which delegation improves governance efficiency and when it amplifies agency costs.

Vote delegation, now common among prominent decentralized governed platforms, creates complex agency conflicts that are critical to understanding the platforms' functionality and growth. Our model builds on the insights in Abadi and Brunnermeier (2024); Sockin and Xiong (2023); Cong et al. (2022) (see the literature review in Section 2). In comparison to prior work that studies how incentive schemes can be optimally embedded in token design, we take the token design as given, and instead derive

implications for the optimal design of monitoring and compensation within token-governed platforms.

The remainder of the paper is organized as follows. Section 2 and Section 3 present descriptive evidence on vote delegation in token-financed platforms and review the related literature. Section 4 presents the model, and Section 5 derives the main results. Section 6 reports simulation implementation and results. Section 7 discusses implications for governance design. Section 8 concludes.

2 Institutional Details and Related Literature

2.1 An Overview of Tokens' Utility

A startup that builds a digital platform can raise external financing by issuing *utility tokens*. These usage-linked claims resemble forward contracts on platform access, i.e., tokens allow holders to redeem them for digital services once the platform becomes operational. Utility tokens do not confer cash-flow rights. Early investors therefore hold tokens largely for speculative purposes, as documented empirically by Fahlenbrach and Frattaroli (2021). This speculative trading and high profit potential have significantly contributed to the rapid growth of digital platforms that are initially financed and later governed through tokens (Howell et al., 2020; Gan et al., 2021).

The early finance and economics literature analyzes the optimal design of such tokens for fundraising. For instance, Gryglewicz et al. (2021) argue that token financing can dominate equity for most early-stage platforms when the token design appropriately combines utility features with cash-flow rights. Cong et al. (2022) highlight the importance of managing token supply dynamically, i.e., issuing too few tokens can deter adoption, whereas issuing too many reduces users' marginal incentives to contribute to platform growth. Prat et al. (2025) further show that optimal supply must balance speculative demand at launch with the convenience yield generated by mature platform usage, implying that supply elasticity is central to sustaining fundamental value.

To improve credibility and reduce agency frictions, many token-financed platforms have replaced or

upgraded their initial utility tokens with *governance tokens*, which may grant holders (future) access to the platform but are primarily used to give them decision rights over operating policies and revenue allocation. For example, platforms that earn revenue from providing services to their users may allow governance token holders to vote on whether revenues should be retained, reinvested, or used to repurchase outstanding tokens (similar to buyback programs).

A prominent example is Uniswap, an automated market-making platform (see Lehar and Parlour (2025) for economic analyses), which received early development funding through a grant from the Ethereum Foundation. After the platform matured, Uniswap introduced its governance token, *UNI*, which was partly allocated to early platform users. *UNI* tokens primarily confer control rights, allowing its holders to vote on the platform's fee structure, trading services, and allocation of accumulated revenues.

Other examples include Aave and Compound, which are leading DeFi lending platforms (see Lehar and Parlour (2022); Rivera et al. (2023) for economic analysis), that have issued governance tokens (*AAVE* and *COMP*) with voting rights over the platforms' lending services, including collateral admissibility, risk management, and payout policies. Unlike Uniswap, Aave financed its early development by selling utility tokens (*LEND*) to early investors, which was later converted into the governance token *AAVE*. Compound, by contrast, relied primarily on conventional equity financing and introduced *COMP* only after the platform was operational, using it as a governance and incentive instrument.

As with *UNI*, neither *AAVE* nor *COMP* provides exclusive access rights. Most decentralized platform services are accessible to all participants. Governance tokens can, however, be used within their respective platforms as ordinary assets. For instance, *UNI* may be supplied as liquidity in Uniswap's market-making mechanism, generating fee income for its providers. Similarly, *AAVE* and *COMP* may be supplied as loanable funds on their lending platforms to earn interest. Nonetheless, such activity is rather limited (see Cornelli et al. (2025) for an empirical analysis of decentralized lending platforms). The primary economic function of governance tokens is to allocate control rights without cash-flow rights, rather than to serve as productive inputs to the platform's services.

2.2 An Overview of Token-Based Governance Effectiveness

At a fundamental level, token-based governance systems operate on the premise that governance token holders, who are typically the platform's users, can propose and vote on policy decisions. Greater token ownership (i.e., skin in the game) provides greater voting power and influence. This structure aligns the incentives of users and platform owners and can be potentially welfare-enhancing under certain conditions.

For instance, Abadi and Brunnermeier (2024) argue that effective token-based governance depends on the presence of robust tokenomic design and platform credibility. Issuing hybrid tokens, which combine access to services with claims on platform profits, can mitigate rent extraction by platform owners and enhance efficiency. However, if users suspect that the platform may issue more tokens or divert profits in the future, they will discount the tokens' value, which undermines the benefits of token-based governance. Therefore, it is essential that the platform commits credibly not to dilute the token supply, in order to maintain trust and preserve the effectiveness of governance.

Sockin and Xiong (2023) show that token-based governance can serve as an effective commitment mechanism to decentralize control and prevent platform-user conflicts, but only under conditions where control remains with users and the platform has strong demand. When tokens primarily attract speculative investors, control rights may become concentrated among non-users, increasing the risk of governance outcomes that enable rent extraction from the user base. Consequently, effective token design should incentivize active participation in governance by users to preserve incentive alignment and reduce the likelihood.

In practice, however, governance tokens are designed to convey only control rights (see subsection 2.1). Uniswap's governance framework is among the first tokens that includes a "fee switch" that, if activated, would redirect a portion of trading fees to token holders. To date, this mechanism remains inactive, largely due to concerns that explicit revenue-sharing could trigger classification of the UNI token as a security under the U.S. Howey Test. In addition, since governance engagement itself requires substantial information-acquisition and monitoring costs, it often exceeds the expected private benefits from voting.

As a result, voter participation tends to be low and concentrated, and governance outcomes are dominated by a small number of large, often non-user, token holders (see Appel and Grennan (2023) for early empirical evidence, and Han et al. (2025) for a literature review). Moreover, Rossello (2024) provides evidence that governance token holders strategically participate in voting. Cong et al. (2025) shows that participants in token-based governance platforms take advantage of informed trading opportunities around voting events.

In an effort to improve efficiency, token-financed platforms have experimented with alternative institutional designs. One such design is vote delegation, which has emerged as one of the most persistent governance mechanisms to date. In the chapter that follows, we collect data and present descriptive statistics on the mechanics of vote delegation as a governance design.

3 The Market for Vote Delegations in Token-financed Platforms

The platform Aragon, which provides an institutional infrastructure for token-based platform governance, was among the first to experiment and implement a programmable, on-chain version of *liquid democracy*. It is a well-studied conceptual model in political theory (see, e.g., Blum and Zuber, 2016), with a few experimental studies and real-world applications, in which voters may either vote directly or delegate their voting rights to agents, with the option to revise the delegation at any time.

The programmable, on-chain version of this concept has added the benefit that token holders do not lose their ownership when they delegate their tokens. Technically, governance tokens remain in the holder's wallet, while designated delegates receive only the associated voting power. Delegation is therefore reversible and state-contingent, i.e., when a token holder sells or transfers governance tokens (for example, by depositing them to the platform to earn interest), the delegate's voting power mechanically declines with the holder's reduced balance. This design aims to mitigate participation failures by allowing informed or active agents to act on behalf of passive holders, while ensuring that ultimate control remains with those who retain the underlying governance tokens.

Because of this novel institutional design, there is little empirical work documenting basic patterns of vote delegation on blockchains. To provide initial evidence, we collect transaction-level data on delegation choices across token-based governance systems on Ethereum, Avalanche, Arbitrum, Optimism, Polygon, and Base, and identify 31 platforms with active delegation mechanisms (see Table A.1 for a summary of the statistics of our sample and a description of the platform services). Each token-based governance system constitutes a distinct organizational entity in which participants may either vote directly on platform policies (by self-delegating and retaining voting power) or delegate their voting power to specialized agents. These delegations records resemble standard financial transfers. But instead of, for example, agent A transferring a specific amount of cryptocurrency or tokens to agent B, the delegation transactions show voter A assigning voting power (denominated in a specific amount of cryptocurrency or tokens) to delegate B.

From these novel data, we construct time series of accumulated voting power for each delegate. In addition, the public registry at tally.xyz often links delegate addresses to identifiable profiles (names, descriptions, and social-media accounts), which allowed us to map governance behavior to specific individuals or organizations.

3.1 Delegates and Incentives to become a Delegate

We identify three broad categories of delegates (see Figure A.1 and Table A.2 in the appendix for illustrative examples of identified delegates and their delegated voting power over time). First, we observe the emergence of professional delegates, including specialized advisory firms, investment and trading firms, and technical development teams. Advisory firms actively engage in platform governance, compete strategically to build reputation and influence, and seek to establish themselves as market leaders in governance consultancy. This dynamic is analogous to how startups overinvest under strategic competition (Inderst and Mueller, 2009). Investment and trading firms may participate to exploit informational advantages arising from governance activity. Technical development teams contribute directly to improving platform infrastructure and are often compensated for these contributions, which

gives them strong incentives to participate in governance. These professional actors interact frequently, and maintain ongoing relationships with other influential stakeholders in the entities, including platform founders and core developers.

Second, a large group of delegates consists of individual participants and non-profit organizations, including university-affiliated student groups. Their participation is primarily motivated by learning, experimentation, and engagement rather than financial returns. This group represents the largest number of addresses but participates only sporadically and with limited influence on aggregate outcomes.

Third, we observe a small number of application-based delegates that would like to incentivize voting participation through the use of their platform. For instance, voters can receive financial rewards on these platforms. These application-based delegates currently receive negligible delegated voting power and have little impact on governance outcomes. However, they represent an emerging institutional form that may become more economically relevant as governance mechanisms evolve.

Delegates Compensation Only a limited number of token-based platform governance compensate voting delegates through explicit monetary or performance-based incentives. For instance, Uniswap recently introduced a delegate compensation scheme that pays up to $6,000 per month in UNI tokens to delegates who satisfy predefined participation and engagement criteria.[1]

Similarly, Optimism has implemented a participation-based remuneration program, awarding active delegates 6,000 tokens (approximately $5,000 at the time of distribution) conditional on voting activity.[2] In 2025, Polygon distributed quarterly compensation ranging from $624 to $10,519 to fourteen highly active delegates, with payments denominated in the platform's native token.[3]

[1] See gov.uniswap.org.
[2] See gov.optimism.io.
[3] See polygon.technology.

3.2 Token-based Platform Governance

Token-based platform governance typically begins with open discussion in public forums (e.g., Telegram or Discord), where users and delegates exchange information and debate proposals. Many platforms then conduct a nonbinding preliminary off-chain vote to gauge support and filter out proposals with little backing. Proposals that attract sufficient support advance to a binding vote on the blockchain. These formal on-chain proposals may bundle multiple previously endorsed off-chain items to reduce coordination costs. As such, an on-chain proposal typically includes a written description and summary of what the community is voting on, as well as executable code that will be deployed on the blockchain if the proposal passes.

Formal voting periods are often short and may require either majority approval and/or meeting a participation quorum. Approved proposals are typically subject to a predetermined implementation delay (a "timelock"), which serves as a commitment and risk-mitigation device by giving stakeholders and market participants time to react. After this delay, the approved proposal's code is executed automatically, making the outcome binding by construction. These formal on-chain proposals can modify governance rules, authorize treasury spending, adjust platform policies, or change token incentive structures.

Misconduct by Delegates in an Observable yet Anonymous Governance Environment This token-based platform governance is open to any market participant and transparent, i.e., the platform's software on the blockchain is publicly observable, and key policy parameters can be modified at any time through stakeholder voting. If a proposal is approved, participants can contribute new program modules and integrate them into the platform's core infrastructure. Crucially, stakeholders can also change the rules of the rule-making process, i.e., they can revise admin rights and/or the conditions under which proposals are accepted. This decentralized mode of production is analogous to open, community-driven projects such as Linux or Wikipedia, where continuous contributions allow the system to evolve over time. Unlike Linux or Wikipedia, however, market participants receive governance rights through tradable tokens. Thus, they have control over strategic decisions such as pricing, fees, user incentives, and most

importantly, the allocation of the platform's collective treasury. This transparency can increase credibility and attract participation in pseudonymous environments, but it also increases the risk of opportunistic manipulation.

Surprisingly, we find only 15 incidents in which previously unaffiliated outsiders attempted to exploit governance by passing a proposal whose effective outcome was to redirect treasury funds to an external recipient account (see Table A.3 for more details in the appendix).[4] Across these 15 incidents, the attack pattern is consistent in that the attacker could not take over the community and immediately drain the treasury funds in one step. Instead, they first introduced proposals that altered control privileges or the voting and approval thresholds required for future proposals within the platform's settings. Only after weakening these safeguards did the attackers attempt a second-stage proposal to extract treasury assets. 10 of these 15 incidents succeeded. Successful attacks primarily occurred in environments with low attention, where inactive users allowed repeated trial-and-error attempts without triggering scrutiny. When communities were active, users functioned as an informal security layer, detecting suspicious proposals and warning others publicly. They also disciplined the platform economically by exiting (selling governance tokens), raising the cost and reducing the payoff of the attack (see Figure A.2 for price dynamics during a governance exploit).

Platform's Treasury A platform's treasury refers to the set of assets owned and controlled by token-based platform governance and therefore available for budget decisions (e.g., grants, contributor compensation, incentive programs, or portfolio rebalancing, see, for instance, compound.woof.software/treasury for a breakdown of Compound's treasury holdings).

The assets in the treasury differ from the total value locked (TVL), which measures the market value of assets supplied by users to the platform (see defillama.com for Compound's TVL). TVL is economically closer to customer balances or assets under management, i.e., these assets remain beneficially owned by platform users and are subject to user withdrawal rights under the platform's rules. As a result,

[4]For a more comprehensive catalog of token-based platform exploits, including non-governance cases, see Feichtinger et al. (2024) and defillama.com.

governance cannot treat TVL as a budget item.

By contrast, treasury holdings represent platform-owned reserves that governance can reallocate. In practice, platforms often maintain separate accounts with different mandates. For example, Compound's treasury is spread across multiple accounts, including program-specific budgets, incentive allocations for delegates and users, managed asset portfolios, and accumulated platform revenues generated by the platform's lending and borrowing services.

These treasuries frequently hold a large inventory of the platform's own governance token, which sometimes constitutes the majority of the treasury's value. The main reason for this is operational: many platforms grant funds, rewards and compensation in their native token, and governance may also choose to buy back governance tokens from the market. Consequently, treasury portfolios tend to be less diversified. At the same time, holding a governance token makes the treasury's resources highly sensitive to the token price, effectively making it a leveraged bet on the platform's future adoption and productivity.

4 Model

Building on the descriptive evidence on vote delegation and governance outcomes in Section 3, we introduce a continuous-time, infinite-horizon framework that endogenizes reputation, monitoring, and delegate incentives. The economy is populated by two types of agents: *users* (principals) and *delegates* (agents). All agents are risk-neutral, discount time at rate r, and are subject to liquidity shocks which follow a Poisson process and occur at rate δ.

Users exercise the governance of the platform by delegating their votes at discrete governance proposal events, which occur according to a Poisson process with counter J_t and rate λ. At each governance proposal, users elect a *leading delegate* to govern the platform until the next proposal, and set their compensation over the current mandate (until the next governance vote).

4.1 Delegates

The model features two rotating delegates. At any given instant, a delegate i occupies the role of the leading delegate, or *incumbent*, while the other delegate j is the *challenger*. When hit by a liquidity shock, a delegate liquidates their position and is immediately replaced.

A liquidity shock to the incumbent triggers a rotation of power, such that either the newcomer or the previous challenger takes the lead in the next governance vote. On the other hand, a shock to the challenger merely resets their specific identity without changing the leader.

For now we assume that delegates control a single identity at a time, a behaviour we later show to be optimal in this setting.

Adverse Selection Each delegate has a type unobservable to users: With probability $1 - p_0$, a delegate is **Honest** (H). This type is a standard rational agent who maximizes expected utility from platform compensation and token appreciation. With complementary probability p_0, a delegate is **Malicious** (M). This type is also a rational utility maximizer but has the ability to "attack" the platform, for example by draining the treasury. If a governance proposal occurs at time t, delegate ℓ chooses to attack with probability $q_\ell(\mathbf{S}_t)$, where $\mathbf{S}_t$ is the state variable of the model.

Agents (tokenholders and delegates) hold a belief $p_{\ell,t}$ representing the probability that delegate ℓ is Malicious. As governance attacks can be implemented right after a governance vote, agents update their beliefs only at these discrete events. Notice also that governance votes reveal information only about the *leading delegate*, as the challenger takes no action.

The belief update mechanism is the standard recursive application of Bayes rule: After proposal k, the posterior odds of the delegate being malicious are simply the prior odds scaled by the probability of a malicious delegate not attacking the platform:

$$\frac{p_{\ell,k}}{1-p_{\ell,k}} = \frac{p_{\ell,k-1}}{1-p_{\ell,k-1}} \cdot (1 - q_\ell(\mathbf{S}_{\tau_k})).$$

By iteratively applying the above update rule, the posterior belief after K successful proposals is given by:

$$p_{k,K} = \frac{p_0 \cdot \prod_{k=1}^{K}(1 - q_\ell(\mathbf{S}\tau_k))}{p_0 \cdot \prod_{k=1}^{K}(1 - q_\ell(\mathbf{S}\tau_k)) + 1 - p_0} \tag{1}$$

The posterior probability that the delegate is malicious decreases with every survived proposal, as long as $q_\ell > 0$.

Moral Hazard Besides deciding on the leading delegate, at each governance proposal, the tokenholders propose a compensation contract (ω, x_d) consisting of a wage flow ω and a governance token "stock option" x_D. The compensation can be renegotiated at every governance proposal.

The purpose of this compensation is to induce the incumbent to exert effort $e(\mathbf{S}_t) \in [0, 1]$ and improve the platform's underlying *quality*, Q_t. This quality drives the service value provided to users and follows the diffusion process:

$$dQ_t = \mu_Q(e_t)dt + \sigma_Q dW_t + \gamma_Q(\mathbf{S}_{t^-})dJ_t \tag{2}$$

Where $\mu_Q(e)$ is a concave function representing the expected improvement in platform features, security, or throughput resulting from delegate effort. $\sigma_A dW_t$ is the Brownian volatility term. $\gamma_Q(\mathbf{S}_{t^-})dJ_t$ is .... [Michele: JUMP DOWN ON ATTACK]

Moral hazard enters the model since the realization of the productivity level A_t is observable to all agents, but the specific effort e_t is not due to the volatility.

4.2 Treasury and Tokenomics

Treasury In order to pay the delegates, the platform must hold a treasury and set a token supply policy. We assume the dynamics of the treasury assets on the platform, T_t, follows a jump-diffusion process:

$$dT_t = (rT_t + fn_t - \omega)\,dt + \sigma_T T_t dW_t + \gamma_T(\mathbf{S}_{t^-})dJ_t \tag{3}$$

The drift term captures risk-free interest rate r, net of the delegate's wage ω, plus fee revenues collected from the tokenholders from using the platform. The jump term $\gamma(\mathbf{S}_{t^-})$ accounts for jumps in the governance events: either a total loss from a malicious attack ($\gamma = -T_{t^-}$) or a partial outflow from a departing delegate liquidating their stake. In the latter case, the treasury contracts by the delegate's proportional ownership, $\gamma = -T_{t^-} x_D / X_t$.

Token Supply Dynamics The platform starts with an initial token base X_0. The total supply X_t evolves stochastically, adjusting at each governance proposal (arrival time τ_k) to match the newly approved compensation level x_{D,τ_k}. Let $x_{D,\tau_{k-1}}$ denote the delegate's stake prior to the current vote. The change in the total token supply is the net difference between the new token compensation and the previous one. The supply adjustment is thus is given by:

$$dX_t = (x_{D,t} - x_{D,t^-})dJ_t, \qquad X_{\tau_k} = X_{\tau_{k-1}} + (x_{D,\tau_k} - x_{D,\tau_{k-1}})... \tag{4}$$

Token supply expands when tokenholders vote to increase the delegate compensation and contracts in the converse case.

4.3 Tokenholders

Participation Dynamics The number of tokenholders of the platform, n_t, follows a birth-death process. New tokenholders arrive at a Poisson rate $\alpha(A_t)$, which is a concave function of the productivity of the platform:

$$\alpha(A_t) = a_0 A_t^{\epsilon}. \tag{5}$$

Tokenholders choose whether to hold the tokens relative to their private outside option u drawn from a distribution $F(u)$ and face a departure rate from the platform $\delta_n(S_t) > 0$ defined by :

$$\delta_n(\mathbf{S}_t) = \delta \cdot [1 - F(U_i)] \tag{6}$$

The number of active tokenholders is denoted by n and δ is a constant rate of liquidity shock common for all tokenholders. This implies that a tokenholder's lifetime within the token-based platform is exponentially distributed with mean $1/\delta_n(S_t)$. For a state with n_t tokenholders, the total departure rate from the system is $n_t\delta_n(S_t)$. The mean number of tokenholders at time t is :

$$\bar{n}(\mathbf{S}_t) = \frac{\alpha(\mathbf{S}_t)}{\delta_n(\mathbf{S}_t)} \tag{7}$$

Since the entry and exit dynamics follow a non-homogeneous birth-death process, standard results from queueing theory (Eick et al., 1993) establish that, conditional on the state history, the number of tokenholders n_t follows a Poisson distribution, $n_t \sim \text{Pois}(\bar{n}(\mathbf{S}_t))$. This feature allows us to model the voting process as a Poisson Game (Myerson, 1998).

Voting for Platform Governance At each governance proposal, tokenholders must decide whether to delegate their voting power to the current **Incumbent** (i) or the **Challenger** (j). To do so, tokenholders face a costly information acquisition problem. They can choose to pay an attention cost c to become **Informed**, allowing them to identify the **Superior Delegate** (ℓ^*), i.e., the candidate with the highest risk-adjusted value, with certainty. Alternatively, they can remain **Heuristic**, voting based on a noisy signal of delegate quality. We assume heuristic voters delegate to the Superior Delegate ℓ^* with probability $\theta > 1/2$ and to the Inferior Delegate ℓ' with probability $1-\theta$. The superior delegate is defined as the candidate offering the highest continuation value:

$$\ell^* = \underset{\ell \in \{i,j\}}{\arg\max} \left\{ (1 - p_{\ell,t} q_{\ell,t}) U_\ell(\mathbf{S}, x_u) \right\} \tag{8}$$

We denote the inferior delegate as $\ell' = \{i, j\} \setminus \{\ell^*\}$.

Vote Aggregation and Pivotality Since the population follows a Poisson distribution, the votes for the Incumbent (i) and Challenger (j) are independent Poisson variables. Their expected rates, $\bar{n}_i$ and $\bar{n}_j$, depend on whether the Incumbent is currently the Superior Delegate ($\ell^* = i$) or the Inferior Delegate

($\ell^* = j$).

$$\bar{n}_i = (\bar{n}_I + \theta\bar{n}_H) \cdot \mathbf{1}_{\{i=\ell^*\}} + ((1-\theta)\bar{n}_H) \cdot \mathbf{1}_{\{i=\ell'\}} \tag{9}$$

$$\bar{n}_j = (\bar{n}_I + \theta\bar{n}_H) \cdot \mathbf{1}_{\{j=\ell^*\}} + ((1-\theta)\bar{n}_H) \cdot \mathbf{1}_{\{j=\ell'\}} \tag{10}$$

Since the vote counts for each candidate follow independent Poisson distributions, the random variable representing the vote difference $\Delta n_{ij} = n_i - n_j$ follows a Skellam distribution:

$$\mathcal{S}(\Delta n_{ij}; \bar{n}_i, \bar{n}_j) = e^{-(\bar{n}_i + \bar{n}_j)} \left(\frac{\bar{n}_i}{\bar{n}_j}\right)^{\Delta n_{ij}/2} I_{|\Delta n_{ij}|}\left(2\sqrt{\bar{n}_i \bar{n}_j}\right) \tag{11}$$

where $I_{|k|}(\cdot)$ is the modified Bessel function of the first kind.

The incumbent wins if the total weighted vote count is positive. The probability of the incumbent retaining power, denoted $\pi_i(\mathbf{S})$, is:

$$\pi_i(\mathbf{S}) = \mathbb{P}\left(\Delta n_{ij} > \frac{\Delta x_{ji}}{x_u}\right) = \sum_{\Delta n = \lfloor \Delta x_{ji}/x_u \rfloor + 1}^{\infty} \mathcal{S}(\Delta n; \bar{n}_i, \bar{n}_j) \tag{12}$$

The challenger wins with complementary probability $\pi_j(\mathbf{S}) = 1 - \pi_i(\mathbf{S})$.

A single tokenholder is **pivotal** if their vote flips the outcome. The probability of being pivotal is the sum of the Skellam probabilities over the critical range where a single stake x_u bridges the gap:

$$\mathbb{P}_{\text{piv}}(\mathbf{S}) = \sum_{\Delta n = \lfloor \Delta x_{ji}/x_u - 1 \rfloor}^{\lceil \Delta x_{ji}/x_u \rceil} \mathcal{S}(\Delta n; \bar{n}_i, \bar{n}_j) \tag{13}$$

Incentives to Acquire Information Tokenholders acquire information if the expected utility gain from ensuring the Superior Delegate ℓ^* wins exceeds the cost c. The Marginal Benefit of Information (MBI) is derived from the pivotal events where a single informed vote flips the election outcome from the Inferior

Delegate ℓ' to the Superior Delegate ℓ^*:

$$\text{MBI}(\mathbf{S}) = \mathbb{P}_{\text{piv}}(\mathbf{S}) \cdot (1-\theta) \cdot \Big[(1-p_{\ell^*,t}q_{\ell^*,t})U_{\ell^*}(\mathbf{S},x_u) - (1-p_{\ell',t}q_{\ell',t})U_{\ell'}(\mathbf{S},x_u)\Big] \tag{14}$$

4.4 Value Functions

State Space The equilibrium dynamics described above evolve over a state space defined by the quintuplet $\mathbf{S}_t = (A_t, T_t, X_t, p_{i,t}, p_{j,t})$. Here, A_t represents the platform's productivity, T_t the treasury, and X_t the total token supply. The variables $p_{i,t}$ and $p_{j,t}$ denote the reputation scores, specifically, the public belief that the incumbent i (or challenger j) is malicious.

The Tokenholder's Problem The representative tokenholder chooses their token demand x_u and information strategy $\mathcal{I} \in \{0,1\}$ to maximize expected utility. Let $U_i(\mathbf{S}, x_u)$ denote the value function when delegate i is the current incumbent. The Hamilton-Jacobi-Bellman (HJB) equation is:

$$(r+\delta+\lambda)U_i(\mathbf{S},x_u) - \mathcal{L}U_i(\mathbf{S},x_u) = \delta P(\mathbf{S})x_u + \lambda \mathcal{E}_{vote}(\mathbf{S},x_u) \tag{15}$$

The term $\delta P(\mathbf{S})x_u$ is the liquidation value upon a liquidity shock. The term $\mathcal{E}^{vote}$ represents the maximized expected value at the governance event. The tokenholder compares the cost of acquiring information (c) against the benefit of voting for the superior candidate:

$$\mathcal{E}_{vote}(\mathbf{S},x_u) = \max_{\mathcal{I}\in\{0,1\}} \left\{ \sum_{\ell\in\{i,j\}} \pi_\ell(\mathbf{S}|\mathcal{I}) \Big[(1-\hat{q}_\ell)U_\ell(\mathbf{S}^+,x_u) + \hat{q}_\ell U_{-\ell}(\mathbf{S}^-,x_u)\Big] - c\cdot\mathcal{I} \right\} \tag{16}$$

Here, $\hat{q}_\ell = p_{\ell,t}q_{\ell,t}$ is the probability that candidate ℓ is malicious and attacks. The term in the square brackets weights the continuation value between two outcomes: If candidate ℓ governs effectively, the state improves to $\mathbf{S}^+$ and ℓ becomes the new incumbent (U_ℓ). Conversely, if candidate ℓ attacks, the state crashes to $\mathbf{S}^-$ and the challenger candidate ($-\ell$) steps in as the new leader ($U_{-\ell}$).

The Delegate's Problem Delegates maximize the sum of their compensation and capital gains. We distinguish between the **Honest** (H) and **Malicious** (M) types, and their role as Incumbent (i) or Challenger (j). The state vector is $\mathbf{S}_t = (p_{i,t}, p_{j,t}, A_t, T_t, X_t)$. To simplify the exposition, we define two auxiliary payoff functions. The *Liquidation Payoff* represents the value of a delegate who fully exits the platform (e.g., after a liquidity shock):

$$\Omega(\mathbf{S}, x_D) \equiv x_D \left(P(\mathbf{S}) + \frac{T_t}{X_t} \right) \tag{17}$$

The *Follower Payoff* represents the continuation value of a delegate who loses power due to a governance vote and remains in the ecosystem as a challenger:

$$\Psi_k(\mathbf{S}, x_D) \equiv P(\mathbf{S}) x_D + V_j^k(\mathbf{S}, 0), \quad k \in \{H, M\} \tag{18}$$

The Honest Incumbent The honest incumbent chooses effort e to maximize value. The HJB equation separates the continuous dynamics from the expected value jump at the governance event:

$$(r+\lambda) V_i^H(\mathbf{S}, x_D) - \mathcal{L} V_i^H(\mathbf{S}, x_D) = \max_{e \geq 0} \left\{ \omega - e + \lambda \left[\mathcal{E}_{gov}^H(\mathbf{S}, x_D) - V_i^H(\mathbf{S}, x_D) \right] \right\} \tag{19}$$

where $\mathcal{E}_{gov}^H$ denotes the expected post-governance value. With probability $\chi = \frac{\delta}{\delta+\lambda}$, the delegate is hit by a liquidity shock. Otherwise, they stand for re-election:

$$\begin{aligned} \mathcal{E}_{gov}^H(\mathbf{S}, x_D) = \chi \Omega(\mathbf{S}, x_D) + (1-\chi) \Big[& \pi_i(\mathbf{S}) V_i^H(\mathbf{S}^+, x_D) \\ & + (1 - \pi_i(\mathbf{S})) \left((1-\bar{q}_j) \Psi_H(\mathbf{S}, x_D) + \bar{q}_j V_i^H(\mathbf{S}^-, x_D) \right) \Big] \end{aligned} \tag{20}$$

The term $\bar{q}_j V_i^H(\mathbf{S}^-)$ captures the case where the challenger wins and attacks (probability $\bar{q}_j$), so the honest delegate is reinstated to manage the damaged platform.

The Malicious Incumbent The malicious incumbent faces a similar problem but possesses the strategic option to attack upon winning a governance vote. Their HJB equation is:

$$(r+\lambda)V_i^M(\mathbf{S},x_D) - \mathcal{L}V_i^M(\mathbf{S},x_D) = \omega + \lambda\left[\mathcal{E}_{gov}^M(\mathbf{S},x_D) - V_i^M(\mathbf{S},x_D)\right] \tag{21}$$

The expected governance value $\mathcal{E}_{gov}^M$ incorporates the optimal attack decision:

$$\begin{aligned}\mathcal{E}_{gov}^M(\mathbf{S},x_D) = \chi\Omega(\mathbf{S},x_D) + (1-\chi)\Big[&\pi_i(\mathbf{S})\max\left\{V_i^M(\mathbf{S}^+,x_D),\ T_t + \Psi_M(\mathbf{S}^-,0)\right\} \\ &+ (1-\pi_i(\mathbf{S}))\Psi_M(\mathbf{S},x_D)\Big]\end{aligned} \tag{22}$$

If the malicious delegate chooses to attack, they seize the treasury T_t and re-enter the collapsed ecosystem $\mathbf{S}^-$ as a follower (Sybil attack[5]), captured by the term $\Psi_M(\mathbf{S}^-,0)$.

4.5 Equilibrium Concept

The equilibrium is determined by the interaction of the optimal policy functions: the delegates' effort e^*, the malicious delegates' attack probability q^*, the tokenholders' information acquisition probability ρ^*, and the resulting aggregate dynamics of the ecosystem.

Delegate Strategies Incumbents of both types exert effort e to improve platform productivity. While their strategic horizons may differ, both types benefit from the token price appreciation induced by higher productivity during their tenure. The optimal effort e_k^* for a delegate of type $k \in \{H, M\}$ balances the marginal cost of exertion against the marginal increase in their continuation value. From the first-order condition of the respective HJB equations, the optimal efforts satisfy:

$$\mu_A'(e_k^*(\mathbf{S}))\frac{\partial V_i^k(\mathbf{S},x_D)}{\partial A} = 1, \quad \text{for } k \in \{H,M\} \tag{23}$$

[5]A type of cyberattack in which a single malicious actor creates numerous fake identities (or accounts) to gain disproportionate influence.

In addition to effort, the malicious incumbent determines the attack probability $q^* \in [0,1]$ to maximize the expected value at the governance event. This choice is the solution to the linear maximization problem:

$$q^*(\mathbf{S}) \in \underset{q \in [0,1]}{\arg\max} \Big\{ q \cdot \Big[T_t + \Psi_M(\mathbf{S}^-, 0) \Big] + (1-q) \cdot V_i^M(\mathbf{S}^+, x_D) \Big\} \tag{24}$$

This linear structure implies a "bang-bang" solution driven by the sign of the net benefit of attacking, denoted by $\Delta(\mathbf{S})$.

Stability Refinement To ensure the malicious strategy is robust to small perturbations ("trembling hand"), we refine the equilibrium concept. Consider an ϵ-perturbed game where the attack probability is constrained to $q \in [\epsilon, 1-\epsilon]$. Let $\{q(\mathbf{S}; \epsilon_n)\}_{n=1}^{\infty}$ be a sequence of equilibrium strategies corresponding to a sequence $\epsilon_n \to 0$. The strategy $q^*(\mathbf{S})$ is **stable** if this sequence converges to q^* in the L^1-norm weighted by the equilibrium invariant distribution of states, $\Phi(\mathbf{S})$:

$$\lim_{n \to \infty} \int_{\mathcal{S}} |q(\mathbf{S}; \epsilon_n) - q^*(\mathbf{S})| \, d\Phi(\mathbf{S}) = 0 \tag{25}$$

Tokenholder Strategies Tokenholders face a collective action problem. The equilibrium information acquisition probability $\rho^* \in [0,1]$ solves the fixed-point problem where the Marginal Benefit of Information (MBI) equals the cost c, subject to boundary constraints:

$$\rho^*(\mathbf{S}) \in \underset{\rho \in [0,1]}{\arg\max} \Big\{ \rho \cdot \Big[\text{MBI}(\mathbf{S}, \rho_{-i}) - c \Big] \Big\} \tag{26}$$

Market Clearing The equilibrium price $P(\mathbf{S})$ equates the fixed token supply to the aggregate demand. Given the delegate's holding x_D, the market clearing condition is:

$$X_t = x_D + x_u^*(\mathbf{S}, P(\mathbf{S})) \tag{27}$$

where x_u^* satisfies the tokenholder's first-order condition $\partial U_i / \partial x_u = P(\mathbf{S})$.

Definition 1: Stable Markov Perfect Equilibrium A Stable Markov Perfect Equilibrium (MPE) consists of a set of value functions for the tokenholders $\{U_i, U_j\}$, the incumbents $\{V_i^H, V_i^M\}$, and the challengers $\{V_j^H, V_j^M\}$; a set of policy functions $\{e^*, q^*, \rho^*, x_u^*\}$; a price function $P(\mathbf{S})$; and a stationary probability measure $\Phi(\mathbf{S})$ defined over the state space $\mathcal{S}$, such that:

1. **Tokenholder Optimality:** Given the price $P(\mathbf{S})$ and beliefs $p_{\ell,t}$, the demand x_u^* and information strategy ρ^* solve the HJB equation (Eq 30) and the maximization problem in Eq (37).

2. **Delegate Optimality:**

   - The honest incumbent's effort e^* satisfies the FOC in Eq (34).
   - The malicious incumbent's effort e^* and attack probability q^* solve the HJB equation (Eq 35) and the maximization in Eq (35).

3. **Market Clearing:** The price function $P(\mathbf{S})$ ensures the token market clears at every state $\mathbf{S} \in \mathcal{S}$ (Eq 39).

4. **Belief Consistency:** The reputation processes $p_{\ell,t}$ evolve according to Bayes' rule (Eq 1), consistent with the equilibrium attack probability $q^*(\mathbf{S})$.

5. **Consistent State Evolution:** The state vector $\mathbf{S}_t$ evolves according to the aggregate law of motion induced by the optimal strategies $\{e^*, q^*, x_{D,t}\}$. Specifically, the drift and diffusion of A_t are determined by e^*, and the jump intensities for treasury T_t and beliefs are determined by q^*.

6. **Distributional Stability:**

   - The measure $\Phi(\mathbf{S})$ is the invariant distribution of the state process $\mathbf{S}_t$ under the optimal strategies.
   - The malicious delegate's strategy q^* satisfies the stability criterion in Eq (36) with respect to $\Phi(\mathbf{S})$.

5 Analysis

We characterize the equilibrium of the model defined in Section 4 by analyzing the interplay between tokenholder oversight and delegate incentives. We begin by examining the delegates' problem, showing that anonymity precludes screening based on inferred effort. Because malicious delegates hold a valuable option to expropriate the treasury, they are incentivized to exert effort levels exceeding those of honest types to secure incumbency. We argue that while the anonymity of the decentralized environment renders the preservation of identity impossible, a hypothetical authenticated regime could induce effort through screening provided agents are sufficiently patient. Next, we turn to the tokenholders, showing that voting incentives are non-linear in the platform's capitalization and strictly decreasing in the size of the user base. Finally, we analyze the aggregate dynamics resulting from these equilibrium policies. We show that the interaction between asset accumulation and governance efficacy generates endogenous instability, characterized by cyclical boom-bust dynamics.

5.1 Effort Choice and the Impossibility of Performance-Based Screening

We first analyze the effort incentives of delegates and the resulting implications for governance. A central result of this section is that the decentralized nature of the platform creates a "competence paradox": the agent with the most detrimental long-term intentions (the Malicious type) paradoxically possesses the strongest short-term incentives to improve the platform.

5.1.1 First-Best Benchmark

To establish an efficiency benchmark, we define a social planner who maximizes the total surplus of the token-based platform, unconstrained by the agency frictions present in the decentralized equilibrium. The planner's objective is to maximize the expected present value of treasury inflows net of the real cost of effort. We denote the social welfare function by $\mathcal{V}(\mathbf{S})$.

The planner's value function $\mathcal{V}(\mathbf{S})$ solves the HJB equation:

$$r\mathcal{V}(\mathbf{S}) = \max_{e \geq 0} \{f\bar{n}(\mathbf{S}) - e + \mathcal{L}_e \mathcal{V}(\mathbf{S})\} \tag{28}$$

As before, the first-best level of effort, e^{FB}, is characterized by the first-order condition equating the marginal cost of effort to the marginal social value of productivity:

$$\mu'_A(e^{FB}) \frac{\partial \mathcal{V}(\mathbf{S})}{\partial A} = 1 \tag{29}$$

5.1.2 The Malicious Incentive

We now turn to the incentives of the delegates. While both Honest (H) and Malicious (M) types benefit from the token appreciation associated with a more productive platform, the Malicious type derives an additional, private benefit from platform growth: a larger treasury increases the payoff from a potential future attack.

This creates a misalignment in incentives where the "bad" type is, in fact, more motivated than the "good" type.

Proposition 1: Adverse Incentives Conditioning on a separating strategy where types play their privately optimal best-response, a Malicious delegate exerts strictly greater effort than an Honest delegate:

$$e^*_M > e^*_H \tag{30}$$

The proof, detailed in Appendix B3, relies on showing that the value of the attack option is strictly increasing in the platform's productivity, regardless the compensation delegate compensation scheme.

The Impossibility of Screening This ranking ($e^*_M > e^*_H$) creates a barrier to governance in an anonymous environment. In standard principal-agent settings, the "bad" type typically has a higher marginal cost of effort (lower ability). This allows the principal to implement a *separating equilibrium*

by offering a menu of contracts: a high-wage/high-performance contract selected by the good type, and a low-wage/low-performance contract for the bad type.

In our setting, the Single-Crossing Property holds, but the direction of the incentive is *inverted*. The Malicious type derives *higher* marginal utility from effort than the Honest type because every unit of productivity increase inflates the value of their attack option. Consequently, any incentive scheme (w, x_D) high enough to satisfy the Honest delegate's participation constraint for high effort will be *even more attractive* to the Malicious delegate.

Thus, tokenholders cannot screen for honesty by demanding performance. If they attempt to do so, the Malicious delegate will simply mimic the Honest type (or even outperform them), biding their time until the treasury T_t reaches a critical threshold for an attack. This forces the equilibrium into a *pooling* outcome, where tokenholders must price the risk of malice into the platform's valuation, treating all delegates as potentially hostile.

Authenticated Benchmark To isolate the friction caused by anonymity, we consider a counterfactual "Authenticated" regime. In this setting, delegate identities are permanent and transparent (e.g., linked to real-world legal identities or "Soulbound" tokens). This has two critical implications for the model dynamics:

1. **Permanent Exclusion:** A delegate who attacks the platform or is voted out for malfeasance cannot re-enter the ecosystem. The"Sybil" value of resetting one's reputation drops to zero ($V_{\text{reset}} = 0$).

2. **Legal/Off-chain Recourse:** The platform may be able to impose an exogenous penalty $\psi > 0$ on a malicious actor (e.g., legal action), effectively lowering the net payoff of an attack.

In the anonymous setting, the Malicious delegate's continuation value upon attacking included the ability to return as a challenger: $T_t + \Psi_M(\mathbf{S}^-, 0)$. In the authenticated setting, this continuation value collapses to $T_t - \psi$.

While the Malicious delegate still has a strong incentive to exert effort to inflate T_t prior to an attack,

the permanence of identity allows tokenholders to utilize *intertemporal screening*. By back-loading compensation (e.g., through vesting schedules), tokenholders can leverage the different time horizons of the two types.

Consider a contract offering a high deferred reward $\mathcal{W}_{def}$ conditional on a long period of sustained high effort without an attack.

- The **Honest** type, who never intends to attack, views this as a standard discounted cash flow problem. If their discount rate r is sufficiently low, they accept the contract.
- The **Malicious** type faces a trade-off: to claim the deferred reward $\mathcal{W}_{def}$, they must delay their attack. However, delaying the attack carries the risk that the opportunity might vanish (e.g., liquidity shocks or displacement by a challenger).

Proposition 2: Ranking of Equilibrium Effort In any stable Markov Perfect Equilibrium, the observable effort of a Malicious delegate is indistinguishable from that of an Honest delegate. Both are strictly below the socially optimal level. The full ranking of incentives is:

$$e_H^* \leq e_M < e^{FB} \tag{31}$$

Conversely, in an authenticated environment with permanent identity, there exists a critical discount rate $\bar{r}$ such that if delegates are sufficiently patient ($r < \bar{r}$), a separating equilibrium exists. Tokenholders can implement a high-effort contract that satisfies the Honest type's participation constraint but violates the Malicious type's incentive compatibility constraint. In this case,

$$e_H^* < e_M < e^{FB} \tag{32}$$

The proof is provided in Appendix B2. This ranking highlights the dual inefficiency of the decentralized equilibrium: effort is sub-optimal relative to the first-best (due to agency costs), yet screening is impossible because the adversarial type has the highest private marginal benefit of exertion.

This result recovers a classic insight from the dynamic contracting literature (see, e.g. DeMarzo and Sannikov, 2006; Biais et al., 2007; DeMarzo and Fishman, 2007; He, 2012; Gryglewicz et al., 2020; He et al., 2017; Winton and Yerramilli, 2021). For instance, when effort is unobservable and agents face short-term incentives tied to future rents, a malicious agent may optimally exert higher effort than an honest one in order to build reputational capital and expand the value of opportunistic deviations . When identity is preserved, long-term reputation mechanisms can substitute for short-term monitoring. It highlights that the "impossibility of screening" derived in Section 5 is not a result of the technology's complexity, but specifically of the *anonymity* constraint which removes the ability to punish via exclusion.

5.2 Contract Design

The adverse selection problem of unobservable delegate types creates an information rent for the malicious delegate. It can always mimic an honest type to build reputation, giving them a strategic advantage. An honest type, by definition, cannot mimic a malicious one. Consequently, the principal, the tokenholders from the platform, must design a contract that satisfies the incentive-compatibility constraint of the malicious type to prevent attacks. This contract, which is just sufficient to secure the good behaviour of the malicious type, is inefficiently designed from the perspective of the honest type. The honest delegate is thus penalized by the existence of malicious types, receiving a compensation package that is lower than what could be offered in a world of perfect information.

The principal's core challenge is that a token-based contract, while incentivizing effort that boosts productivity and grows the treasury, simultaneously increases the value of the prize from an attack. The optimal contract therefore operates on a frontier, balancing the marginal benefit of higher productivity against the marginal cost of elevated attack risk.

The delegate's time preference is the key determinant of the optimal contract structure. Tokens are a long-duration asset whose value is highly sensitive to the discount rate, r. For a patient delegate (low r), tokens represent a highly cost-effective instrument for satisfying the no-attack constraint, as a small grant can generate a large increase in their continuation value. This intuition is formalized below.

Proposition 3: Primacy of Tokens for Patient Agents There exists a threshold for the discount rate, $\bar{r} > 0$, such that for any delegate with a rate of time preference $r < \bar{r}$, the cost-minimizing, incentive-compatible contract is dominated by token-based compensation (x_D).

The proof of Proposition 7 can be found in Appendix B7.

Proposition 4: No-Attack Equilibrium A no-attack equilibrium ($q^*(\mathbf{S}) = 0$) can be sustained if the delegate's continuation value is sufficiently high. This is achieved if: (a) the delegate's token compensation (x_D) is sufficiently large, and (b) their expected tenure is sufficiently long (i.e., the exit shock rate, δ, is sufficiently low).

5.3 Voting Incentives and Apathy

The efficacy of platform governance rests on the collective vigilance of tokenholders. While the platform is decentralized, the information required to distinguish a Superior Delegate (ℓ^*) from an Inferior one (ℓ') is costly. The incentives to participate in governance are driven by the Marginal Benefit of Information (MBI), which is the product of the probability of being pivotal and the utility gain from ensuring the better candidate wins. Our analysis reveals that these incentives are highly sensitive to the scale of the platform.

The Scale Effect and the Paradox of Growth In a Poisson voting game, the probability of any single tokenholder being pivotal, $\mathbb{P}_{\text{piv}}$, is the likelihood that their single vote (or stake x_u) tips the balance of the election. As the platform successfully attracts more users, the expected number of voters $\bar{n}(\mathbf{S}_t)$ increases.

Following the properties of the Skellam distribution, as $\bar{n}$ grows large, the probability of being pivotal decays at a rate of approximately $1/\sqrt{\bar{n}}$. This leads to a rational apathy, i.e., as the platform becomes more successful in terms of user adoption, the individual incentive to remain informed vanishes. When $\bar{n}$ is high, the "dilution of pivotality" ensures that even a large personal stake feels insignificant relative to the aggregate voting pool. Consequently, in the limit of mass adoption, tokenholders cease to pay the cost c and revert to heuristic-based voting.

Heuristics and Information Substitution The parameter θ represents the accuracy of the "noisy signal" available to heuristic (uninformed) voters. We find that θ and active participation ρ^* act as strategic substitutes. If wisdom of the crowd is already high (high θ), the incremental value of becoming informed is low. In contrast, if the public signal is highly unreliable ($\theta \to 1/2$), the value of certain information peaks, provided that the user base is small enough to maintain pivotality.

Proposition 5: Governance Participation In any Stable Markov Perfect Equilibrium, the probability of information acquisition $\rho^*(\mathbf{S})$ exhibits the following comparative statics:

1. **Scale Decay:** ρ^* is strictly decreasing in the expected user base $\bar{n}$. There exists a threshold $\bar{N}$ such that for all $\bar{n} > \bar{N}$, $\rho^* = 0$, and governance relies entirely on heuristics.
2. **Heuristic Crowding-Out:** ρ^* is strictly decreasing in the precision of heuristic signals θ. High-quality public signals reduce the incentive for private investigation.
3. **Vulnerability Scaling:** The Marginal Benefit of Information (MBI) is increasing in the "prize" of an attack T_t, but this is often offset by the Scale Decay, creating a window of vulnerability where a large treasury is protected by an increasingly apathetic population.

The narrative of the model thus suggests that DAOs are most secure when they are "small and hungry". As they scale toward the first-best productivity levels, the very growth that creates value for the treasury simultaneously erodes the collective oversight necessary to protect it.

5.4 Attack Incentives

While tokenholders succumb to apathy, the Malicious incumbent actively monitors the state $\mathbf{S}_t$ to optimize the timing of an attack. The attack probability $q^*(\mathbf{S})$ is the result of an indifference condition between the certain "booty" of the treasury and the discounted value of future incumbency.

The "Prize" and the "Golden Goose" The strategic calculus is primarily driven by the ratio of the treasury to productivity. The treasury T_t represents the "prize" of the attack, while productivity A_t

represents the health of the "golden goose" that provides future rewards. As the treasury grows relative to the delegate's stake, the incentive to exit via expropriation dominates.

Reputational Feedback The model features a "reputation trap". As public suspicion $p_{i,t}$ grows, the delegate's expected future tenure shrinks. This reduction in the "shadow of the future" makes the immediate payoff of an attack more attractive. Conversely, the presence of a highly suspicious challenger j acts as a shield for the protocol, as it increases the incumbent's job security and, by extension, their incentive to remain honest.

Proposition 6: Comparative Statics of Attack In any Stable Markov Perfect Equilibrium, the equilibrium probability of an attack $q^*(\mathbf{S})$ satisfies:

1. **Expropriation Incentive:** q^* is strictly increasing in the treasury size T_t and the total token supply X_t (due to dilution of the delegate's share).
2. **Fundamental Stability:** q^* is strictly decreasing in the platform's productivity A_t.
3. **Tenure Effect:** q^* is increasing in the incumbent's perceived malice $p_{i,t}$ and decreasing in the challenger's perceived malice $p_{j,t}$.

This characterization implies that the "window of vulnerability" for a token-based platform is specifically when a period of high growth has inflated the treasury (T_t), but a subsequent slowdown in productivity (A_t) or a rise in suspicion ($p_{i,t}$) reduces the delegate's long-term stake in the system.

5.5 Cyclicity and Boom-Bust Dynamics

The interaction between the delegate's effort incentives and the tokenholders' voting apathy generates a recurring pattern of growth and collapse. This cyclical behaviour suggests that instability is not an external shock but an endogenous feature of decentralized governance.

The system is governed by a two-way feedback loop. First, higher productivity A_t increases the platform's value and the treasury T_t, which according to Proposition 5, triggers a decay in voting vigilance ρ^*. Second, as oversight weakens and the treasury grows, the malicious delegate's incentive to attack q^* increases (Proposition 6).

When the platform is too successful, it becomes a victim of its own scale. The prize for malfeasance is maximized exactly when the cost of detection is minimized by rational apathy. This leads to the following formal results:

Proposition 7: Stability and Ergodicity If the drift of the state vector $\mu(\mathbf{S})$ is globally mean-reverting, then the stochastic process for the state $\mathbf{S}_t$ is ergodic and admits a unique stationary distribution, $\pi(\mathbf{S})$.

Despite the existence of a long-run distribution, the platform rarely rests at its mean. Instead, the strength of the feedback between governance quality and asset accumulation leads to persistent cycles.

Proposition 8: Endogenous Oscillations When the cross-feedback effects between productivity (A_t) and governance quality are sufficiently strong relative to their direct mean-reversion, the linearized generator of the system possesses complex conjugate eigenvalues. This implies that the platform exhibits cyclical boom-bust dynamics around its stationary mean.

6 Model Simulation

6.1 Implementation and Main Results

We simulate the model using a baseline parametrization in which the key parameters are calibrated using real-world values from the DeFi market and our data, as described in Appendix C, which also details the numerical implementation, reports the full set of parameter values, and presents all simulation figures. The simulations reveal pronounced cyclical dynamics in both the treasury and token price, consistent with the theoretical predictions of Proposition 6. Treasury growth, and, by implication, price appreciation, arises endogenously from sustained increases in protocol productivity, illustrated

in the first panel of Model Simulation. Productivity growth in the simulations is partly driven by strictly positive effort exerted by the leading malicious delegate. To assess the effect of moral hazard of the malicious delegate, we compute a counterfactual effort path corresponding to an honest leading delegate holding fixed the realized state variables. This counterfactual effort is uniformly lower than the effort chosen by a malicious leader, in line with the comparative-statics result in Proposition 2. The simulation relies on some deliberate simplifications and functional-form choices, all detailed in Appendix C. In particular, productivity growth, effort costs, and treasury inflows are specified using smooth and monotone functional forms that admit interior solutions and stable numerical behaviours. These choices are designed to isolate the interaction between effort provision, governance risk, and treasury dynamics in a transparent and tractable way.

Model Simulation graphs about here

The final five panels are the most informative, as they illustrate the microfoundations underlying the cyclical dynamics in the platform's development. At the onset of a delegate's tenure, reputation is low, implying that tokenholders face a high Marginal Benefit of Information (MBI) acquisition, MBI, which exceeds the cost c_A of acquiring information (monitoring). As a result, information acquisition is widespread, and the equilibrium mass of informed tokenholders satisfies $nI^*(\mathbf{S_t}) \simeq \bar{n}(A_t)$, generating the high and noisy plateaus observed in the corresponding panel. During this phase, intensive monitoring disciplines the leading malicious delegate, resulting in a low equilibrium attack probability $q_i^*(\mathbf{S_t})$. Over time, conditional on the absence of attacks and the provision of relatively high effort, the malicious delegate's reputation gradually improves, as reflected in the path of $p_{i,t}$. As beliefs converge toward high reputation, the marginal benefit of monitoring, proportional to $(1 - p_{it}q_i^*(S_t))$, declines and eventually falls below the information-acquisition cost c_A. At this point, monitoring ceases in equilibrium, and the economy transitions to a regime in which no tokenholder acquire information. Contingent with this regime switch, the attack probability $q_i^*(S_t)$ rises sharply, reflecting the absence of effective oversight. Once monitoring collapses, a malicious delegate finds it optimal to exploit the next governance opportunity by attacking the treasury. This attack results in a discrete contraction in

treasury holdings T_t and therefore a sharp decline in the governance token price P_t. Depending on how tokenholders form their expectations this attack can either be completely unanticipated or partly anticipated, we plot both price paths for each possibility (how those different price functions are coded is detailed in Appendix C. Following the attack, the incumbent is replaced, beliefs reset, and a new low-reputation delegate is installed. Monitoring again becomes optimal, and the equilibrium reverts to the high-information regime with $nI^*(\mathbf{S_t}) \simeq \bar{n}(A_t)$.

Consistent with the model's predictions, belief updating from low to high reputation endogenously weakens tokenholders' incentives to monitor the leading delegate, reflecting a free-riding problem in information acquisition once the delegate is perceived as sufficiently trustworthy. Only once monitoring fully collapses does the malicious delegate gain the opportunity to attack the platform. The simulations help visualize how macro-level governance cycles emerge from micro-level dynamic incentives. As an additional **robustness check**, we run 50 independent simulations to assess whether the baseline simulation reported in the main analysis is representative rather than idiosyncratic. For each variable of interest, we compute the mean trajectory across simulation paths and the associated confidence intervals. This exercise confirms that the observed dynamics are systematically generated by the model. The qualitative patterns remain unchanged: cyclical behaviour persists across simulations. As expected, cycles are attenuated in the averaged paths, reflecting the smoothing induced by aggregation and the fact that attacks occur at different times across realizations. Importantly, the confidence intervals remain sufficiently tight, indicating that the dynamics are not driven by simulation noise.

6.2 Delegate's Contract Design to Reduce the Frequency of Attacks

We next examine how delegate contract design affects governance risk by varying the composition of delegate compensation between fixed (dollar-denominated) pay and token-based rewards. Using repeated simulations, we study how changes in token payouts influence the frequency of governance attacks, holding all other parameters constant. This exercise allows us to isolate the incentive effects of token-based compensation and assess whether increased reliance on token payments can discipline

delegate behaviour.

Delegate token compensation versus Number of attacks graph about here

The graph plots the mean number of attacks per path as a function of the delegate's token-based compensation. Consistent with Proposition 3, increasing the share of compensation paid in governance tokens is associated with a decline in the frequency of attacks. The effect of token-based compensation on attack frequency is the product of two countervailing forces. On the one hand, when a delegate is compensated primarily in tokens, an attack that drains the treasury or undermines platform fundamentals induces a sharp decline in the token price. Because the delegate internalizes this price impact through their own compensation, the expected profitability of an attack is reduced, dampening ex ante incentives to deviate. On the other hand, higher token-based compensation necessarily requires greater token issuance, which mechanically expands the circulating supply and depresses the token price even absent attacks. Relative to a fixed dollar-denominated base pay, this dilution effect lowers the real value of delegate compensation and weakens the delegate's effective stake in the platform. When dilution dominates incentive alignment, the disciplining role of token-based compensation is eroded, and attack incentives may persist. The observed flattening and non-monotonicity in attack frequency therefore reflect the interaction of these two forces. Our results indicate that token-based compensation can dampen governance risk, but only when contract design balances incentive alignment against dilution; large token emissions alone are insufficient to eliminate attacks.

7 Implications for Governance Design

Our paper yields several normative implications for the efficiency of the design of decentralized governance platforms. First, delegate compensation must be carefully designed ex ante so that honest behaviour remains optimal even when the ex-ante probability of a successful attack is high. In the model, insufficiently strong incentive allows malicious delegates to exert positive effort, build reputation, and delay attacks until monitoring incentives collapse. As shown in the simulations, without those incentives,

delegates accumulate trust endogenously through the tokenholders' Bayesian updating, escape oversight, and eventually exploit governance opportunities. While token-based compensation can, in principle, align delegate incentives by increasing their exposure to the token price, aggressive token issuance may have counter-intuitive effects. Large token emissions mechanically dilute token value, reducing the real compensation of delegates relative to fixed dollar-denominated pay and weakening the extent to which delegates internalize the cost of an attack. When this dilution effect dominates, increasing token-based compensation does not meaningfully reduce attack incentives and may even leave them unchanged. The simulations therefore demonstrate that large token emissions are not a panacea for reducing governance risk: without careful calibration, higher nominal token payouts can fail to discipline malicious behaviour and may inadvertently preserve the very incentives they are intended to mitigate.

Second, we highlight the importance of the information environment in which governance takes place. Platforms characterized by an excessive frequency of low-quality or uninformative proposals accelerate Bayesian updating corresponding to environments in which the marginal benefit of monitoring is reduced more rapidly, pushing tokenholders into a free-riding equilibria. By contrast, a more selective and informative proposal flow slows belief convergence, sustains monitoring incentives, and delays the transition to the unmonitored regime in which attacks become likely. Our paper suggests that governance design should explicitly manage proposal frequency and informativeness, rather than treating delegates participation intensity as unambiguously reputation-enhancing.

8 Conclusion

In a standard principal-agent setting, principals mitigate moral hazard not only by strengthening external enforcement (monitoring) but also by informal discipline (trust) based on the agent's reputation and perceived integrity. However, a central problem is that trust and thus agent's effective commitment is not fixed, but endogenous. In other words, an agent responds to the expected private returns from opportunism and to the governance environment itself.

We formalize this mechanism in a model of token-based platform governance and derive a reputation-monitoring cycle that can impede platform growth. The sequence is as follows. An agent, who may be malicious that has initially low reputation, supplies high effort, thereby raising platform value and accumulating reputational capital. As the marginal benefit of acquiring information and monitoring a highly reputational agent falls, so does the monitoring intensity from the principal (the tokenholders). Once oversight is low, the agent exploits residual control rights embedded in the delegated governance tokens to launch an attack (e.g., draining the treasury). The attack triggers a sharp collapse in platform value. Following the attack, the incumbent is replaced, beliefs reset, a new low-reputation delegate is instated and the cycle restarts.

This dynamic agency conflict arises from the interaction of adverse selection (the principal cannot perfectly screen agents ex ante) and moral hazard (effort and intent are imperfectly contractible ex post). This conflict is not idiosyncratic to token-based platform governance. Rather, it is a general feature of environments in which performance-relevant effort and constraints are difficult to verify. In such settings, reputation can substitute for formal enforcement until it becomes too effective. In anonymous governance environments, however, we show that the long-run costs of reputational loss are muted, as agents can re-enter with reset identities thus weakening intertemporal discipline and amplifies the cyclical dynamics generated by reputation building and subsequent opportunistic behaviour.

One implication of the model is that malicious agents may rationally front-load their efforts. By overinvesting in their initial performance, these agents inflate the value of the platform and their own credibility. This accelerates the endogenous withdrawal of monitoring and maximizes the profits from a later attack. Honest agents lack the option value of sabotage and therefore have weaker incentives to mimic this extreme early effort. Consequently, mitigating this cycle of governance failure requires mechanisms that align the agent's continuation value with platform value, even after reputation is established. For example, the platform can implement token-based delegates' compensation and bonding schemes which are state-contingent and forfeitable. In this setting, this could mean requiring meaningful token stakes with lockups and slashing, rather than relying on fixed cash payments or wages, so that

the agent better internalizes downside risk. In addition, platforms can manage the flow of governance proposals to avoid sudden, belief-driven reductions in oversight, i.e., premature trust based on cheap, activity-based signals. Platforms can also avoid artificial reputation building by conditioning reputational metrics on costly commitments (e.g., agent could supply a collateral as insurance if an attack happens) rather than on raw governance activity.

References

Abadi, J. and Brunnermeier, M. (2024). Token-based platform governance. *Journal of Financial Economics*, 162:103951.

Aghion, P., Algan, Y., Cahuc, P., and Shleifer, A. (2010). Regulation and distrust. *Quarterly Journal of Economics*, 125(3):1015–1049.

Appel, I. and Grennan, J. (2023). Control of decentralized autonomous organizations. In *AEA Papers and Proceedings*, volume 113, pages 182–185. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

Biais, B., Mariotti, T., Plantin, G., and Rochet, J.-C. (2007). Dynamic security design: Convergence to continuous time and asset pricing implications. *Review of Economic Studies*, 74(2):345–390.

Blum, C. and Zuber, C. I. (2016). Liquid democracy: Potentials, problems, and perspectives. *Journal of Political Philosophy*, 24(2):162–182.

Catalini, C., de Gortari, A., and Shah, N. (2022). Some simple economics of stablecoins. *Annual Review of Financial Economics*, 14(1):117–135.

Cong, L. W., Li, Y., and Wang, N. (2022). Token-based platform finance. *Journal of Financial Economics*, 144(3):972–991.

Cong, L. W., Rabetti, D., Wang, C. C., and Yan, Y. (2025). Centralized governance in decentralized organizations. Working paper.

Cornelli, G., Gambacorta, L., Garratt, R., and Reghezza, A. (2025). Why DeFi lending? Evidence from Aave V2. *Journal of Financial Intermediation*, page 101166.

DeMarzo, P. M. and Fishman, M. J. (2007). Optimal long-term financial contracting. *Review of Financial Studies*, 20(6):2079–2128.

DeMarzo, P. M. and Sannikov, Y. (2006). Optimal security design and dynamic capital structure in a continuous-time agency model. *Journal of Finance*, 61(6):2681–2724.

Diamond, D. W. (1984). Financial intermediation and delegated monitoring. *Review of Economic Studies*, 51(3):393–414.

DuPont, Q. (2017). Experiments in algorithmic governance: A history and ethnography of "The DAO," a failed decentralized autonomous organization. In *Bitcoin and beyond*, pages 157–177. Routledge.

Eick, S. G., Massey, W. A., and Whitt, W. (1993). Mt/g/ queues with sinusoidal arrival rates. *Management Science*, 39(2):241–252.

Fahlenbrach, R. and Frattaroli, M. (2021). ICO investors. *Financial Markets and Portfolio Management*, 35(1):1–59.

Feichtinger, R., Fritsch, R., Heimbach, L., Vonlanthen, Y., and Wattenhofer, R. (2024). SoK: Attacks on DAOs. Working paper.

Ferreira, D. and Li, J. (2024). Governance and management of autonomous organizations. Working paper.

Ferreira, D., Li, J., and Nikolowa, R. (2023). Corporate capture of blockchain governance. *Review of Financial Studies*, 36(4):1364–1407.

Gan, J., Tsoukalas, G., and Netessine, S. (2021). Initial coin offerings, speculation, and asset tokenization. *Management Science*, 67(2):914–931.

Gryglewicz, S., Mayer, S., and Morellec, E. (2020). Agency conflicts and short-versus long-termism in corporate policies. *Journal of Financial Economics*, 136(3):718–742.

Gryglewicz, S., Mayer, S., and Morellec, E. (2021). Optimal financing with tokens. *Journal of Financial Economics*, 142(3):1038–1067.

Han, J., Lee, J., and Li, T. (2025). A review of DAO governance: Recent literature and emerging trends. *Journal of Corporate Finance*, page 102734.

Harvey, C. R., Hasbrouck, J., and Saleh, F. (2024). The evolution of decentralized exchange: Risks, benefits, and oversight. Working paper.

Harvey, C. R., Ramachandran, A., and Santoro, J. (2021). *DeFi and the Future of Finance*. John Wiley & Sons.

Harvey, C. R., Saleh, F., and Sverchkov, R. (2025). An economic model of the L1-L2 interaction. Working paper.

He, Z. (2012). Dynamic compensation contracts with private savings. *Review of Financial Studies*, 25(5):1494–1549.

He, Z., Wei, B., Yu, J., and Gao, F. (2017). Optimal long-term contracting with learning. *Review of Financial Studies*, 30(6):2006–2065.

Howell, S. T., Niessner, M., and Yermack, D. (2020). Initial coin offerings: Financing growth with cryptocurrency token sales. *Review of Financial Studies*, 33(9):3925–3974.

Inderst, R. and Mueller, H. M. (2009). Early-stage financing and firm growth in new industries. *Journal of Financial Economics*, 93(2):276–291.

John, K., Kogan, L., and Saleh, F. (2023). Smart contracts and decentralized finance. *Annual Review of Financial Economics*, 15(1):523–542.

Lehar, A. and Parlour, C. (2025). Decentralized exchange: The Uniswap automated market maker. *Journal of Finance*, 80(1):321–374.

Lehar, A. and Parlour, C. A. (2022). Systemic fragility in decentralized markets. Working paper.

Makarov, I. and Schoar, A. (2022). Cryptocurrencies and decentralized finance (DeFi). *Brookings Papers on Economic Activity*, 2022(1):141–215.

Myerson, R. B. (1998). Extended poisson games and the condorcet jury theorem. *Games and Economic Behavior*, 25(1):111–131.

Prat, J., Danos, V., and Marcassa, S. (2025). Fundamental pricing of utility tokens. *Management Science*.

Rivera, T. J., Saleh, F., and Vandeweyer, Q. (2023). Equilibrium in a DeFi lending market. Working paper.

Rossello, R. (2024). Blockholders and strategic voting in DAOs' governance. Working paper.

Sockin, M. and Xiong, W. (2023). Decentralization through tokenization. *Journal of Finance*, 78(1):247–299.

Winton, A. and Yerramilli, V. (2021). Monitoring in originate-to-distribute lending: Reputation versus skin in the game. *Review of Financial Studies*, 34(12):5886–5932.

Appendix A: Motivating Evidence

Table A.1. Our Sample of Token-Based Platform Governances

This table reports our sample of 31 token-financed platforms with delegated governance and the highest observed levels of vote delegation. The "Blockchain" column identifies the ledger on which voting and delegation are recorded. The final column reports the total supply of governance tokens. Because these tokens are traded on exchanges, the outstanding supply is distributed across platform treasuries, the circulating free float, and holdings controlled by tokenholders.

Token-Based Platform	#Proposals	#Tokenholders	#Delegates	Blockchain	#Governance Tokens
Compound	459	218,889	18,510	Ethereum	10,000,000
Aave	401	118,655	151,276	Ethereum	16,000,000
PoolTogether	117	8,546	1,444	Ethereum	10,000,000
InstaDapp	111	10,766	545	Ethereum	100,000,000
Gitcoin	110	92,025	20,905	Ethereum	100,000,000
Fei	97	13,503	3,618	Ethereum	1,000,000,000
Uniswap	84	383,815	48,986	Ethereum	1,000,000,000
Arbitrum	82	1,849,401	445,014	Arbitrum	9,999,998,978
Unlock	67	4,089	515	Ethereum	1,000,875
Reflexer Ungovernor	61	3,774	78	Ethereum	957,122
ENS	55	66,610	37,445	Ethereum	100,000,000
Ampleforth	54	26,020	1,080	Ethereum	15,297,897
Seamless Protocol	48	105,818	2,548	Base	100,000,000
Yam Finance	40	11,899	2,135	Ethereum	15,164,603
Cryptex	37	3,037	101	Ethereum	10,000,000
Inverse	29	4,053	784	Ethereum	719,000
Babylon Finance	28	1,128	109	Ethereum	1,000,000
Indexed	27	5,374	924	Ethereum	10,000,000
Radworks	27	7,748	185	Ethereum	99,998,580
Threshold Network	27	8,801	135	Ethereum	11,155,000,000
Optimism	25	1,379,570	305,663	Optimism	4,294,963,292
Pooh	25	20,324	67	Ethereum	420,690,000,000,000
Idle DAO	24	3,703	150	Ethereum	13,000,000
DEGENX Ecosystem	21	524	118	Avalanche	20,680,502
BIGCAP	19	206	44	Ethereum	100,000,000
Hop	17	12,905	9,494	Ethereum	1,000,000,000
Hifi DAO	15	6,252	341	Ethereum	166,968,917
DIMO	13	79,799	307	Polygon	684,498,958
Anvil	11	3,004	2,726	Ethereum	100,000,000,000
Rari Capital	10	22,909	1,843	Ethereum	1,000,000,000
Aggregated Finance	10	1,124	34	Ethereum	628,086,797,657

Notes: These platforms provide a menu of digital financial services that mirrors familiar functions of the financial system, but implemented through smart contracts (software) rather than a intermediary (Harvey et al., 2021; Makarov and Schoar, 2022; John et al., 2023). The dominant platform products are credit-market services, i.e., overcollateralized borrowing and lending with algorithmic collateral requirements and liquidation rules (see, e.g., Compound, Aave, Seamless Protocol, Inverse, Hifi DAO, Rari Capital). Uniswap provides exchange services (Lehar and Parlour, 2025). The platforms Fei, Ampleforth and Reflexer Ungovernor manage the issuance and stabilization of stablecoins (Catalini et al., 2022). PoolTogether offers savings accounts. Babylon Finance, Indexed, Idle DAO, DEGENX Ecosystem, and Aggregated Finance resemble asset managers. Both Arbitrum and Optimism operate L2 blockchains (Harvey et al., 2025). Another set of platforms provides services to the blockchain infrastructure that lowers trading, settlement, and other frictions across vanues (e.g., InstaDapp, Anvil, Hop, Threshold Network). ENS and Unlock provide a digital identity service. Gitcoin and Radworks focus fundraising and grant allocation. Finally, Pooh is best characterized as a speculative token whose value is primarily driven by social media.

Table A.2. An Overview of Delegates

This table lists the top delegates by delegated voting power in our sample of 31 DAOs with the largest total delegations. We classify delegates into four organizational types. First, registered for-profit companies that offer services and use delegation and governance engagement as a signalling device. Second, student-run organizations that use governance engagement as a networking and learning opportunity. Third, applications, which are platform businesses that aim to improve governance by facilitating user participation through their products and services. Fourth, individuals, whose participation may be driven by private benefits, reputational concerns, or intrinsic motivation to contribute to the ecosystem.

Name	Delegate in	Summary
A. Companies		
Andreessen Horowitz (a16z)	Compound, Fei, Rari Capital, Uniswap	Andreessen Horowitz is a venture capital firm headquartered in Menlo Park, California (*LinkedIn*: founded in 2009, 201-500 employees).
Avantgarde	Arbitrum	Avantgarde Finance is a London-headquartered digital-asset investment manager and advisory firm (*LinkedIn*: founded in 2018, 11-50 employees). It markets itself as a DeFi asset manager that provides treasury management, portfolio construction, risk monitoring, and research/advisory services to DAOs and foundations, and it also offers regulated digital-asset funds for eligible investors.
Blockworks Research	Arbitrum, Uniswap	Blockworks Research is the research arm of Blockworks, a crypto-focused media and data company founded in 2018 (*LinkedIn* reports 2-10 employees, lists New York as headquarters, and lists 2022 as the founding year). It produces institutional research and uses that research capacity to inform governance participation and public commentary on policy choices.
Consensys	Arbitrum, Compound, Hop, Optimism, Uniswap	Founded in 2014, ConsenSys is an infrastructure and software company in the Ethereum ecosystem (*LinkedIn* reports 501-1,000 employees and lists its headquarters in Fort Worth, Texas). Through its security and auditing activities, it has interacted with governance and security-related processes, including discussions around governance structures and safety mechanisms.
Entropy Advisors	Arbitrum	Entropy Advisors is a small governance advisory team that is publicly dedicated to Arbitrum (*LinkedIn*: 2-10 employees). It presents itself as a specialized delegate that invests in proposal analysis and in the design of decision rules, with the goal of improving the quality and credibility of collective choice.
Gauntlet	Aave, Arbitrum, Compound, Fei, Optimism, Rari Capital, Uniswap	Gauntlet was founded in 2018 and is headquartered in New York (*LinkedIn*: 51-200 employees). It is a quantitative risk management and advisory firm that applies economic and statistical modeling to evaluate incentive schemes and risk exposures in decentralized financial markets. The firm produces public recommendations used in governance and parameter setting.
GFX Labs	Aave, Arbitrum, Compound, Hop, Optimism, Uniswap	GFX Labs is a research and product firm that participates in governance across multiple protocols (*LinkedIn* lists Chicago as headquarters and reports 11-50 employees). In Arbitrum, it contributes to policy debates by publishing voting rationales and by focusing on issues such as treasury programs, research initiatives, and institutional design.
Pantera Capital	Ampleforth, Arbitrum, Compound, InstaDapp	Pantera Capital is an alternative investment firm that runs venture and hedge-fund strategies focused on digital-asset markets and related financial infrastructure (*LinkedIn*: founded in 2003, headquartered in Menlo Park, California, 51-200 employees). The firm states that it has invested in this sector since 2013 and positions itself as an early U.S. institutional manager dedicated to this market segment.

Name	Delegate in	Summary
StableLab	Aave, Arbitrum, Hop, InstaDapp, Optimism, Uniswap	StableLab is a professional delegation and governance advisory organization (*LinkedIn*: founded in 2021, 11-50 employees). It provides structured vote analysis and proposals that emphasize incentives, accountability, and long-run budget discipline.
Wintermute	Aave, Arbitrum, Compound, Optimism, Uniswap	Wintermute was founded in 2017 and is headquartered in London (*LinkedIn*: 201-500 employees). The firm operates as a large proprietary trading and liquidity provision firm in DeFi markets. Its business model resembles market making, and it engages in governance discussions as a delegate.
B. Student organizations at Universities or non-profits		
Blockchain at Berkeley	Arbitrum, Compound, ENS, Fei, Rari Capital, Uniswap	Blockchain at Berkeley was founded in 2016, and is one of the earlier large university organizations in this space. Public descriptions emphasize education and building activity, and early reporting indicates it reached roughly one hundred members at that stage of its development (*LinkedIn*: 11-50 members, headquartered in Berkeley, California).
Blockchain at Columbia	Aave, Arbitrum, Compound, ENS, Hop, Optimism, Uniswap	Blockchain at Columbia is a student organization focused on research and education. In July 2021, it announced a reorganization of its governance structure to increase participation beyond current students. Its stated objective is to strengthen continuity by integrating alumni and external contributors into its decision-making and programming (*LinkedIn*: 11-50 members).
Blockchain at UCLA	Aave, Arbitrum, Compound, ENS, Fei, Rari Capital, Optimism, Uniswap	Founded in 2017, Blockchain at UCLA is a university student organization that organizes events and training to lower entry costs for students and help members build practical skills valued by employers in the sector (*LinkedIn*: 51-200 members, headquartered in Los Angeles).
Franklin DAO	Aave, Arbitrum, Compound, Hop, Optimism, Uniswap	Franklin DAO is a student-founded organization that emerged from the University of Pennsylvania community and is associated with the EduDAO initiative. It positions itself as a coordination layer that supports education, community building, and project development through a shared institutional framework (*LinkedIn*: 11-50 members, headquartered in Philadelphia).
Michigan Blockchain	Aave, Arbitrum, Compound, Optimism, Uniswap	Founded in 2018, Michigan Blockchain is a student-run organization at the University of Michigan. It combines education with applied projects, including research and consulting-style work (*LinkedIn*: 51-200 members, headquartered in Ann Arbor).
PGov	Arbitrum, Compound, ENS, Gitcoin, InstaDapp, Optimism, Uniswap	PGov is a small delegation team that provides proposal analyses and voting rationales. They claim to focus on growth, transparency, and diversifying the Arbitrum treasury into more stable allocations.
Stanford Blockchain Club	Aave, Compound, Uniswap	The Stanford Blockchain Club, which traces its origins back to 2014, describes itself as a long-running student organization focused on education and entrepreneurship. Its public materials emphasize programming that connects students with industry professionals and highlight faculty advising to ensure continuity and maintain quality (*LinkedIn*: 2-10 members, headquartered in Stanford, California).

Name	Delegate in	Summary
404Gov (via 404 DAO)	Arbitrum, Optimism, Uniswap	404Gov is the governance and delegation arm of 404 DAO. 404 DAO describes itself as a nonprofit community organization that focuses on education and professional networking, including programs like Web3 ATL and an accelerator initiative aimed at local talent and early-stage teams. In Arbitrum and other DAOs, 404Gov acts as a professional delegate that aggregates delegated voting power, publishes written voting rationales, and participates in governance workstreams such as committees and councils (*LinkedIn*: 2–10 employees, headquartered in Atlanta, Georgia).
C. Applications		
LobbyFi	Arbitrum, Optimism	LobbyFi is an application that intermediates between dispersed token holders and actors who demand voting power. It frames governance rights as a tradable input into policy outcomes and offers a mechanism through which vote holders can monetize participation while buyers aggregate influence for particular proposals.
Mux Protocol	Arbitrum	MUX is a derivatives trading application focused on perpetual futures. It provides a trading venue that aggregates liquidity and supports leveraged positions, and it has been described in public reporting as the continuation of the MCDEX lineage that rebranded and relaunched under the MUX name around 2022.
EventHorizon.vote	Aave, Arbitrum, Uniswap	EventHorizon.vote positions itself as an application layer that aggregates voting power from dispersed holders and organizes participation through a pooled delegate structure (*LinkedIn*: founded in 2022, 2-10 employees).
D. Individuals		
humpy	Compound	humpy is a pseudonymous governance participant listed as a delegate in Compound. Public delegate profiles show cross-DAO participation and voting activity, which fits a model where reputation and repeated interaction can discipline voting behavior even without a formal corporate structure (*Boardroom*).
L2BEAT	Arbitrum, Hop, Optimism, Uniswap	L2BEAT is a public-goods organization that produces monitoring and analytical benchmarks for Ethereum scaling systems. It positions itself as an independent information intermediary that reduces informational frictions for users and governance participants by standardizing risk and performance disclosures.
Lefteris Karapetsas	ENS, Gitcoin, Hop, Optimism, Uniswap,	Lefteris Karapetsas is a delegate who is also publicly associated with rotki, which positions itself as a portfolio and accounting tool, and his governance relevance comes from active participation and delegated voting in these ecosystems (*Tally*).
Linda Xie	Arbitrum, Compound, Gitcoin, Optimism	Linda Xie is a governance participant listed as a delegate in Optimism and Gitcoin. *Tally*, *LinkedIn* describe her as a co-founder of Scalar Capital, an investment firm headquartered in San Francisco that reports a founding year of 2017 and a size of roughly 2-10 employees.
olimpio	Aave, Arbitrum, Compound, DIMO, ENS, Gitcoin, Hop, Optimism, PoolTogether, Seamless Protocol, Uniswap	olimpio is a pseudonymous delegate who participates in governance discussions and presents a focus on Ethereum scaling. The delegate statement emphasizes proposal analysis and cross-ecosystem experience, with the stated goal of improving decision quality in collective governance.

Table A.3. Governance Exploits in Token-Based Platform Governance

This table summarizes incidents in which the open and transparent governance environment of a token-based platform was exploited. We classify the cases as follows: (i) self-dealing by insiders (e.g., founders, core contributors, and large stakeholders), and (ii) external attempts to misuse governance to reallocate treasury assets. Success equals one if treasury resources were transferred against the collective interest. For a more comprehensive list of platform exploits, including incidents not driven by governance, see Feichtinger et al. (2024) and defillama.com.

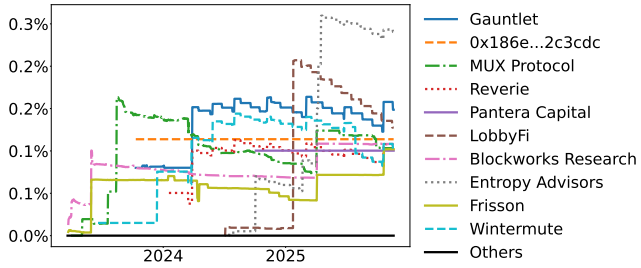
DAO	Timestamp	Summary	Success	Source
Aragon	May 2, 2023	A group "RFV Raiders" accumulated governance tokens to obtain majority voting rights and sought to pass a proposal that would liquidate (or reallocate) treasury assets for private gain.	0	blog.aragon.org
Atlantis Loans	Jun 10, 2023	An attacker obtained approval for a proposal that quietly granted them effective control privileges (Proposal 49), Proposal 52). After the implementation delay (Queue), they exploited their authorizations to transfer approximately ~$1.16 million to an external account.	1	medium.com, Binance
Audius	Jul 24, 2022	An attacker proposed reallocating 18.5 million governance tokens (~$6 million) from the community treasury to an external account (Proposal 84, Proposal 85).	1	rekt.news, Ethereum
Beanstalk	Apr 17, 2022	The attacker temporarily acquired majority voting rights by using short-term borrowing, also known as flash loans, approved Proposal 18 and Proposal 19, and reallocated approximately ~$77 million from the treasury.	1	medium.com, Ethereum
Build Finance	Feb, 11 2022	The attacker passed a proposal that reassigned control privileges, giving them effective authority over governance token issuance and the treasury (Proposal 7). Once in charge, they issued over one million new governance tokens and extracted funds from liquidity reserves (~$470,000).	1	beosin.com, Ethereum, Ethereum, Ethereum
Compound	Dec 2, 2020	Insiders exploited their control privileges to divert approximately ~$12 million from the treasury to external accounts.	1	rekt.news
Compound	Sep 29, 2021	Tokenholders approved Proposal 62 which introduced an implementation error in the rewards mechanism, and some users then claimed extra governance tokens (~$147 million)	1	rekt.news
Compound	May 6, 2024	The account "Humpy" proposed allocating governance tokens into the goldCOMP vault. The proposal was flagged as risky and failed to meet the participation threshold.	0	Proposal 247
Compound	Jul 15, 2024	A renewed proposal to allocate governance tokens to goldCOMP was rejected following public opposition and concerns about governance risk.	0	Proposal 279
Compound	Jul 28, 2024	"Golden Boys" proposed allocating governance tokens (~$25 million) to goldCOMP; the proposal passed on low participation but was cancelled before implementation, so no funds were transferred.	0	Proposal 289
CurioDAO	Mar 23, 2024	An implementation error allowed the attacker to inflate voting rights by issuing ~1 billion governance tokens, receive control privileges, and divert approximately ~$16 million from the treasury.	1	rekt.news, Ethereum
Fortress Loans	May 9, 2022	An attacker passed a harmful rule change and manipulated the platform's reference-price mechanism, enabling them to borrow against overstated collateral values and extract roughly ~$3 million from the lending markets.	1	rekt.news

DAO	Timestamp	Summary	Success	Source
Indexed Finance	Nov 22, 2023	A hostile proposal took control over the treasury's implementation-delay mechanism. Proposal 24 and Proposal 27 passed, but prior to implementation the proposer accepted a side payment of about $10,000 to withdraw the action.	1	xuantify.com
Synthetify	Oct 01, 2023	An attacker submitted 10 proposals (nine largely empty and one hostile, see: Proposal) to transfer ~$230,000 to themselves. They met the participation threshold using their own holdings.	1	blockworks.co
Tornado Cash	May 13, 2023	The attacker copied a previous proposal description but added a backdoor to the underlying code, which would have granted them control of the platform (Proposal 20). After it passed, they gave themselves ~1.2 million governance tokens. The community later regained control, so the takeover was not ultimately successful.	0	decrypt.co, rekt.news

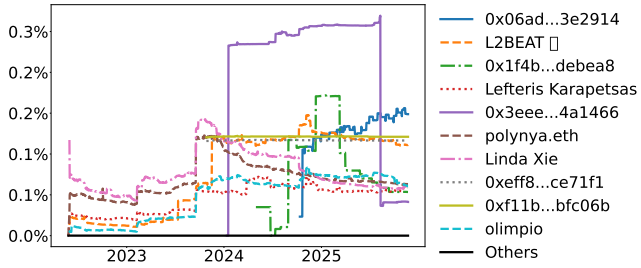
Figure A.1: Evolution of Vote Delegations per DAO

This figure shows how delegated voting power (as a percentage of the total number of tokens) shifts over time within each organization (as of December 2025). It reflects whether control is becoming more concentrated (meaning a small number of actors hold a growing share of the voting rights) or more competitive, with voting power distributed across multiple delegates whose relative positions change over time. Some actors are more transparent than others. For example, in Arbitrum DAO, most delegates are registered companies, whereas in the other DAOs, delegates are mostly individuals. Furthermore, Arbitrum, Optimism, and ENS display a more competitive governance landscape, with leadership among major delegates rotating over time, while Compound and Uniswap have more stable delegate structures.

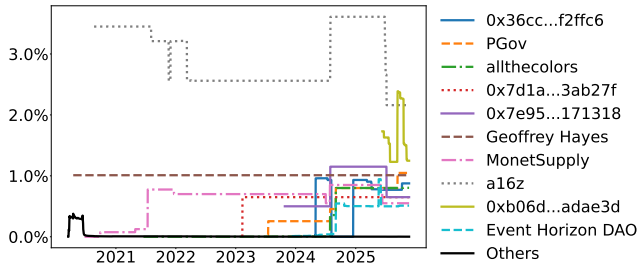
(a) Arbitrum DAO on Arbitrum



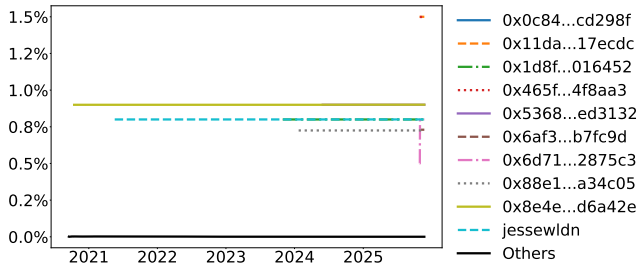
(b) Optimism DAO on Optimism



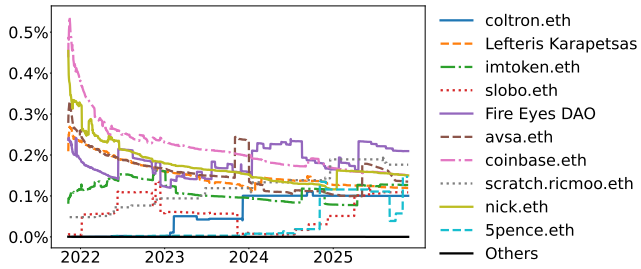
(c) Compound DAO on Ethereum



(d) Uniswap DAO on Ethereum



(e) ENS DAO on Ethereum



(f) Seamless DAO on Base

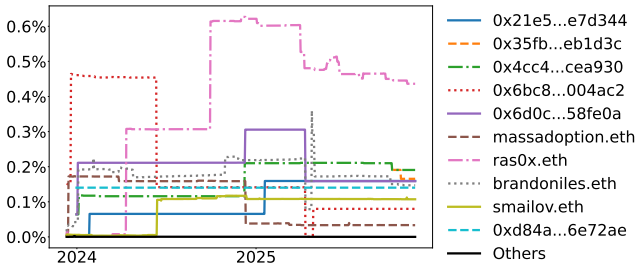
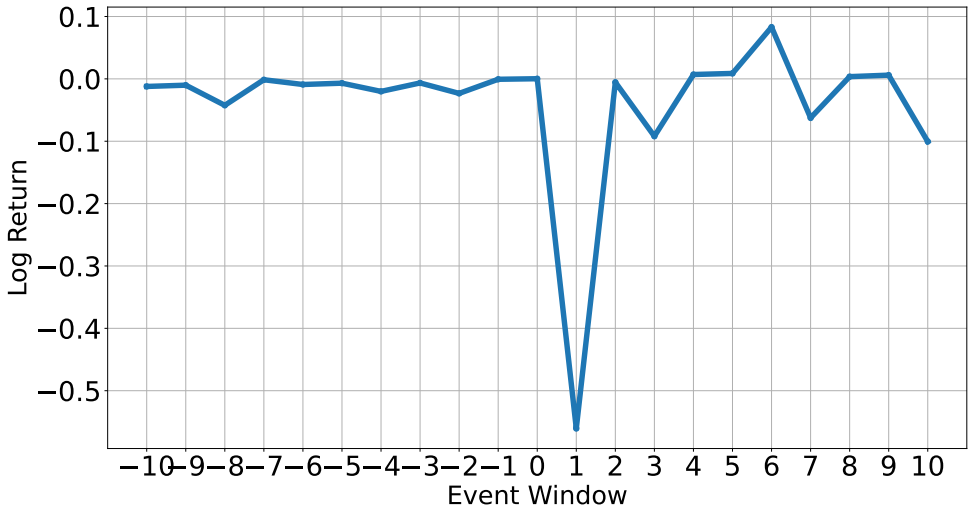


Figure A.2: Token price around a hostile takeover event
This figure plots the token price from ten days before to ten days after the hostile takeover attack in the open access governance environment. Day 0 marks the the attack day (see Table A.3). The event window illustrates the short-run market response to governance security risk and the associated change in expected control rights.



Appendix B: Proof sketches

B2: Proof of Proposition 2

Proof. The optimal effort for any agent (type $k \in \{H, M\}$ or the social planner) is determined by their respective first-order condition, which equates the marginal cost of effort to its marginal benefit. As the cost function is convex ($c'' > 0$), effort is monotonically increasing in the perceived marginal value of productivity. The proposition is established by proving the following ranking of these marginal values:

$$\frac{\partial W_H}{\partial A} \leq \frac{\partial W_M}{\partial A} < \frac{\partial \mathcal{V}}{\partial A} \tag{B1}$$

Step 1: Comparing Delegate Types ($e_H^* \leq e_M^*$). The malicious delegate's value function, W_M, can be decomposed into the honest delegate's value, W_H, plus the value of the option to attack the treasury,

$V_{\text{option}}(\mathbf{S})$.

$$W_M(\mathbf{S}) = W_H(\mathbf{S}) + V_{\text{option}}(\mathbf{S})$$

Higher productivity A_t leads to a larger expected future treasury, making the attack option more valuable. Thus, the marginal value of productivity for the malicious type is strictly greater: $\frac{\partial W_M}{\partial A} = \frac{\partial W_H}{\partial A} + \frac{\partial V_{\text{option}}}{\partial A} > \frac{\partial W_H}{\partial A}$.

This gives the malicious delegate a strategic choice:

- In a **separating equilibrium**, the malicious delegate acts on their natural incentive, choosing their privately optimal effort e^*_M, which is strictly greater than e^*_H.
- In a **pooling equilibrium**, the malicious delegate strategically mimics the honest type to protect their reputation. They deliberately choose to exert the lower effort level, such that $e^*_M = e^*_H$.

Combining these possibilities, it must be that $e^*_M \geq e^*_H$ in any equilibrium.

Step 2: Comparing to First-Best ($e^*_M < e^{FB}$). The social welfare function $\mathcal{V}(\mathbf{S})$ is the sum of the value to all token holders (V_{TH}) and the value to the delegate. Even considering the malicious type, $\mathcal{V}(\mathbf{S}) = V_{TH}(\mathbf{S}) + W_M(\mathbf{S})$. An increase in productivity benefits all other token holders, so $\frac{\partial V_{TH}}{\partial A} > 0$. The delegate, being self-interested, ignores this positive externality. The planner, by contrast, internalizes it. Therefore, the marginal social value of productivity is strictly greater than the malicious delegate's private marginal value:

$$\frac{\partial \mathcal{V}}{\partial A} = \frac{\partial V_{TH}}{\partial A} + \frac{\partial W_M}{\partial A} > \frac{\partial W_M}{\partial A}$$

This proves that the first-best effort level must be strictly higher than the malicious delegate's effort.

Step 3: Authenticated Banchmark. Let V_H^{auth} and V_M^{auth} be the values in the authenticated regime. The principal offers a contract requiring effort e^* and paying a deferred lump sum Ω at time τ. The Honest type accepts if $e^{-r\tau}\Omega - \int_0^\tau e^{-rt}c(e^*)dt \geq 0$. The Malicious type, however, compares waiting

for Ω against attacking immediately at $t = 0$ to seize T_0. If $T_0 > e^{-r\tau}\Omega$, the Malicious type rejects the waiting strategy. Combined, a separating contract exists if:

$$e \int_0^\tau e^{-rt} dt \leq e^{-r\tau}\Omega < T_0 \tag{B2}$$

For sufficiently small r (high patience), the Honest type's valuation of the future reward $e^{-r\tau}\Omega$ approaches Ω, allowing the principal to set Ω high enough to cover effort costs, while keeping the immediate treasury T_0 low enough (or the delay τ long enough) to dissuade the Malicious type. In the anonymous case, this fails because the Malicious type can attack T_0 *and* re-enter to try for Ω later (Sybil attack), breaking the exclusivity of the trade-off.

B3: Proof of Proposition 3

Proof. The optimal effort for a delegate of type $k \in \{H, M\}$ solves the first-order condition $c'(e_k^*) = \mu'_A(e_k^*)\frac{\partial W_k}{\partial A}$. As $c'' > 0$, effort is monotonically increasing in the marginal value of productivity, $\frac{\partial W_k}{\partial A}$. The proof proceeds in two steps: first, we show that the expected future treasury is increasing in current productivity, and second, we show this implies $\frac{\partial W_M}{\partial A} > \frac{\partial W_H}{\partial A}$.

Step 1: Endogenous Link Between Productivity and Treasury. Let the drift of the treasury process be $\mu_T(\mathbf{S}) = \kappa_T(\bar{T} - T_t) + f_0\bar{n}(\mathbf{S})^\beta - \omega$. We show that $\frac{\partial \mu_T}{\partial A_t} > 0$, which implies that $\frac{\partial \mathbb{E}_t[T_{t+k}]}{\partial A_t} > 0$ for any $k > 0$. By the chain rule:

$$\frac{\partial \mu_T}{\partial A_t} = \frac{\partial \mu_T}{\partial \bar{n}} \frac{\partial \bar{n}}{\partial A_t} \tag{B3}$$

The first term is the effect of the user base on treasury inflows:

$$\frac{\partial \mu_T}{\partial \bar{n}} = f_0 \beta \bar{n}^{\beta-1} > 0 \tag{B4}$$

since $f_0, \beta > 0$. The second term is the effect of productivity on the user base, $\bar{n} = \alpha(\mathbf{S})/\delta_n(\mathbf{S})$. Using the quotient rule:

$$\frac{\partial \bar{n}}{\partial A_t} = \frac{(\partial \alpha / \partial A_t)\delta_n - \alpha(\partial \delta_n / \partial A_t)}{\delta_n^2} \tag{B5}$$

We evaluate the derivatives in the numerator. First, the arrival rate $\alpha(\mathbf{S}) = a_0 \frac{A_t}{1+A_t/\bar{A}}$ is strictly increasing in productivity:

$$\frac{\partial \alpha}{\partial A_t} = \frac{a_0}{(1+A_t/\bar{A})^2} > 0$$

Second, the departure rate δ_n decreases with productivity because higher A_t increases token holder utility $U(\mathbf{S})$, making them less likely to leave: $\frac{\partial \delta_n}{\partial A_t} < 0$. The numerator is therefore $(+)\cdot(+) - (+)\cdot(-) > 0$. This confirms that $\frac{\partial \bar{n}}{\partial A_t} > 0$. Combining the terms, we have $\frac{\partial \mu_T}{\partial A_t} = (+)\cdot(+) > 0$. The treasury's drift is strictly increasing in productivity.

Step 2: Comparison of Marginal Values. The value function of a malicious delegate, $W_M(\mathbf{S})$, can be expressed as the value of an honest delegate, $W_H(\mathbf{S})$, plus the value of the option to attack, $V_{\text{option}}(\mathbf{S}) > 0$. Differentiating with respect to productivity A yields:

$$\frac{\partial W_M}{\partial A} = \frac{\partial W_H}{\partial A} + \frac{\partial V_{\text{option}}}{\partial A}$$

The value of the attack option, V_{option}, is an increasing function of the expected future treasury, $\mathbb{E}_t[T_{t+k}]$. As established in Step 1, $\frac{\partial \mathbb{E}_t[T_{t+k}]}{\partial A_t} > 0$. By the chain rule, it follows that $\frac{\partial V_{\text{option}}}{\partial A} > 0$. The existence of this strictly positive malicious incentive implies:

$$\frac{\partial W_M}{\partial A} > \frac{\partial W_H}{\partial A} \tag{B6}$$

Given the first-order condition and $c'' > 0$, a strictly greater marginal benefit of effort implies a strictly greater level of optimal effort. Thus, $e_M^* > e_H^*$. □

Proposition 9: The equilibrium probability of an attack for an incumbent malicious delegate i, $q_i^*(\mathbf{S})$,

has the following comparative statics:

(a) It is increasing in the size of the treasury, T_t.

(b) It is decreasing in the DAO's fundamental productivity, A_t.

(c) It is increasing in the belief that delegate i is malicious, $p_{i,t}$.

(d) It is increasing in the total token supply, X_t.

(e) It is decreasing in the belief that challenger delegate j is malicious, $p_{j,t}$.

B4: Proof of Proposition 4

Proof. The proof relies on the Implicit Function Theorem applied to the malicious delegate's indifference condition, $G(\mathbf{S}, q^*) \equiv W_{\text{in}}(\mathbf{S}; q^*) - V_{\text{attack}}(\mathbf{S}) = 0$. For a stable equilibrium, we assume $\frac{\partial G}{\partial q^*} = \frac{\partial W_{\text{in}}}{\partial q^*} > 0$. The derivative of the equilibrium probability q^* with respect to any state variable S is $\frac{dq^*}{dS} = -(\partial G/\partial S)/(\partial G/\partial q^*)$. The sign of $\frac{dq^*}{dS}$ is therefore the opposite of the sign of $\frac{\partial G}{\partial S}$.

Part (a): Effect of Treasury (T_t). We evaluate $\frac{\partial G}{\partial T_t} = \frac{dW_{\text{in}}}{dT_t} - \frac{\partial V_{\text{attack}}}{\partial T_t}$. Let the delegate's tokens x_d represent a claim on a share $s_d = x_d/X_t < 1$ of the treasury, realizable upon an exit shock δ.

The marginal benefit from an attack is $\frac{\partial V_{\text{attack}}}{\partial T_t} = 1$. The marginal benefit on the continuation value, $\frac{dW_{\text{in}}}{dT_t}$, is the delegate's share multiplied by an effective discount factor, DF, which is strictly less than 1 due to discounting ($r > 0$) and liquidity shocks ($\delta > 0$). Thus, $\frac{\partial V_{\text{attack}}}{\partial T_t} = 1 > s_d > \frac{dW_{\text{in}}}{dT_t}$.[6] Therefore,

$$\frac{dq^*}{dT_t} = -\frac{\partial G/\partial T_t}{\partial G/\partial q^*} = -\frac{(-)}{(+)} > 0$$

[6]The inequality $1 > s_d$ is strict in every non-trivial case.

Part (b): Effect of Productivity (A_t). $\frac{\partial G}{\partial A_t} = \frac{\partial W_{\text{in}}}{\partial A_t} > 0$, as productivity is a fundamental driver of the DAO's long-term value, directly increasing the value of the delegate's stake. Thus,

$$\frac{dq^*}{dA_t} = -\frac{(+)}{(+)} < 0$$

Part (c): Effect of Incumbent's Reputation ($p_{i,t}$). $\frac{\partial G}{\partial p_{i,t}} = \frac{\partial W_{\text{in}}}{\partial p_{i,t}} < 0$, as a higher belief of maliciousness reduces the incumbent's expected tenure and thus lowers continuation value. Thus,

$$\frac{dq^*}{dp_{i,t}} = -\frac{(-)}{(+)} > 0$$

Part (d): Effect of Total Token Supply (X_t). $\frac{\partial G}{\partial X_t} = \frac{\partial W_{\text{in}}}{\partial X_t} < 0$. An increase in the total supply X_t has a dilutionary effect on the equilibrium token price $P(\mathbf{S})$, which lowers the value of the delegate's holdings and their continuation value. Thus,

$$\frac{dq^*}{dX_t} = -\frac{(-)}{(+)} > 0$$

Part (e): Effect of Challenger's Reputation ($p_{j,t}$). $\frac{\partial G}{\partial p_{j,t}} = \frac{\partial W_{\text{in}}}{\partial p_{j,t}} > 0$. A higher belief that the challenger j is malicious ($p_{j,t}$) makes them a less credible electoral threat. This increases the incumbent i's job security and expected tenure, thereby raising their continuation value. Thus,

$$\frac{dq^*}{dp_{j,t}} = -\frac{(+)}{(+)} < 0$$

This completes the proof. □

B5: Proof of Proposition 5

Proof Sketch. The proof relies on constructing a Lyapunov function. A sufficient condition for ergodicity is the existence of a twice-differentiable function $V(\mathbf{S})$, known as a Lyapunov function, such that

$V(\mathbf{S}) \to \infty$ as $\|\mathbf{S}\| \to \infty$, and the expected change of the function is negative outside some compact set. The expected change is given by the infinitesimal generator $\mathcal{L}$ applied to V:

$$\mathcal{L}V(\mathbf{S}) = \nabla V(\mathbf{S}) \cdot \mu(\mathbf{S}) + \frac{1}{2}\text{Tr}\left(\Sigma(\mathbf{S})\Sigma(\mathbf{S})^T \text{Hess}(V(\mathbf{S}))\right)$$

The model's explicit mean-reverting drifts (e.g., $\kappa_A(\bar{A} - A_t)$) ensure that for a suitable choice of $V(\mathbf{S})$ (e.g., a quadratic form), the drift component $\nabla V \cdot \mu$ becomes strongly negative when $\mathbf{S}$ is far from its mean. This term will dominate the diffusion component, ensuring that $\mathcal{L}V(\mathbf{S}) < 0$ for large $\|\mathbf{S}\|$. This negative drift in the Lyapunov function guarantees that the process is positive recurrent and thus converges to a unique stationary distribution. □

B6: Proof of Proposition 6

Proof Sketch. Local dynamics around the stationary mean, $\bar{\mathbf{S}}$, are governed by the eigenvalues of the linearized generator, $\mathcal{L}$. Complex eigenvalues imply oscillatory dynamics.

The eigenvalues of $\mathcal{L}$ are approximated by those of the Jacobian of the drift vector, $J_\mu = \nabla\mu(\bar{\mathbf{S}})$. The economic feedback loop of this cycle dictates the signs of the key off-diagonal elements for the (A, n_I) subsystem:

$$\frac{\partial \mu_A}{\partial n_I} > 0 \quad \text{and} \quad \frac{\partial \mu_{n_I}}{\partial A} < 0$$

The eigenvalues of J_μ are complex if the discriminant of its characteristic equation is negative, i.e.,

$$(\text{Tr}(J_\mu))^2 - 4\det(J_\mu) < 0.$$

The determinant is given by:

$$\det(J_\mu) = \frac{\partial \mu_A}{\partial A}\frac{\partial \mu_{n_I}}{\partial n_I} - \frac{\partial \mu_A}{\partial n_I}\frac{\partial \mu_{n_I}}{\partial A} \tag{B7}$$

Given the signs of the on-diagonal (negative) and off-diagonal terms, the determinant is $(-)(-) - (+)(-) > 0$. Since $\det(J_\mu)$ is positive, the condition for complex eigenvalues can be met if the magnitude

of the cross-effects (which determine the determinant) is sufficiently large relative to the direct dampening effects (which determine the trace). This yields a complex conjugate pair of eigenvalues and proves the existence of local cyclical dynamics. □

B7: Proof of Proposition 7

Proof. The principal's problem is to choose (ω, x_D) to minimize cost $C(\omega, x_D)$ subject to the delegate's incentive-compatibility (IC) constraint: $W_{\text{in}}(\omega, x_D; r) \geq V_{\text{attack}}$.

The delegate's continuation value, W_{in}, is composed of the present value of the wage stream, $PV_\omega = \frac{\omega}{r+\delta}$, and the present value of the token stake, PV_x. The token stake is a long-duration asset, as its value depends on the entire future path of the DAO, while the wage is a perpetuity.

The value of a long-duration asset is more sensitive to the discount rate r than that of a perpetuity. Formally, the elasticity of the value with respect to the discount rate is greater for the token component:

$$\left| \frac{\partial PV_x}{\partial r} \frac{r}{PV_x} \right| > \left| \frac{\partial PV_\omega}{\partial r} \frac{r}{PV_\omega} \right|$$

This implies that as the delegate becomes more patient ($r \to 0$), the value of the token stake, PV_x, increases disproportionately compared to the value of the wage stream.

Therefore, for a sufficiently low r, the marginal increase in W_{in} per unit of cost to the DAO is greater for a token grant x_D than for a wage increase ω. Tokens become the more cost-effective instrument for satisfying the IC constraint. □

B8: Proof of Proposition 8

Proof. A no-attack equilibrium requires the strict inequality $W_{\text{in}}(\mathbf{S}) > V_{\text{attack}}(\mathbf{S})$ for all $\mathbf{S}$. The attack payoff V_{attack} is independent of the contract term x_D and the exit shock δ.

(a) The continuation value W_{in} is strictly increasing in the delegate's token stake x_D. Therefore, for

any given state, there exists a sufficiently large x_D such that the inequality holds.

(b) The delegate's value function solves the HJB equation $(r+\delta)W_{\text{in}} = \dots$. By the implicit function theorem, $\frac{\partial W_{\text{in}}}{\partial \delta} < 0$, meaning a lower exit risk increases the continuation value. Therefore, for any given state, there exists a sufficiently low $\delta > 0$ that raises W_{in} enough to satisfy the inequality.

Combining these, a large enough token stake and a low enough exit risk can make the opportunity cost of an attack prohibitively high, thus deterring it completely. □

Appendix C: Model simulation

Implementation of the simulation

This appendix documents the numerical implementation of the model and the simulation procedures. The goal is to provide a transparent account of the algorithmic structure, functional dependencies, and numerical approximations underlying the simulation, pricing, and welfare calculations.

State variables, timing, and discretization

Time is discretized on a uniform grid with step size Δt, over a finite horizon $[0, \mathcal{T}]$. At each grid point $t_k = k\Delta t$, the economy is summarized by the state vector

$$S_t \equiv (A_t, T_t, X_t, p_{i,t}, p_{j,t}),$$

where A_t denotes protocol productivity, T_t the treasury balance, X_t the total token supply, and $p_{i,t}$, $p_{j,t}$ the posterior probabilities that the incumbent and challenger delegates are honest types (reputations), respectively.

Two delegates i (incumbent/leader) and j (challenger) hold fixed token stakes x_i and x_j. Outside

tokenholders absorb the residual supply

$$x_u(S_t) \equiv \max\{X_t - x_i - x_j, 0\},$$

which clears the token market at each date and enters all pricing and utility calculations.

Law of motion for continuous state variables

Productivity Productivity evolves according to an Euler discretization of a continuous-time diffusion:

$$A_{t+\Delta t} = A_t + \mu_A(e_{\text{lead},t})\,\Delta t + \sigma_A\sqrt{\Delta t}\,\varepsilon_{t+\Delta t}, \qquad \varepsilon_{t+\Delta t} \sim \mathcal{N}(0,1).$$

The productivity drift is specified as a power function of the leading delegate's effort. In particular, the code implements

$$\mu_A(e_{\text{lead}}) = \mu_{0,A} + \mu_{1,A}\, e_{\text{lead}}^{\alpha_A},$$

where $\mu_{0,A}$ is a baseline drift term, $\mu_{1,A} > 0$ scales the marginal effectiveness of effort, and $\alpha_A > 0$ governs the curvature of effort's contribution to productivity growth. The derivative used in the leading delegate's effort first-order condition is therefore

$$\mu_A'(e_{\text{lead}}) = \mu_{1,A}\alpha_A\, e_{\text{lead}}^{\alpha_A - 1}.$$

To avoid numerical instability at or near zero effort when $\alpha_A < 1$, the implementation applies a safeguard: if $e_{\text{lead}} \le 0$, the derivative is evaluated at a small positive value $e_{\min} = 10^{-8}$ rather than at zero.

Treasury The treasury follows

$$T_{t+\Delta t} = T_t + \Big(r_T T_t + f(\bar{n}(A_t)) - \omega\Big)\Delta t + \sigma_T T_t\sqrt{\Delta t}\,\eta_{t+\Delta t}, \qquad \eta_{t+\Delta t} \sim \mathcal{N}(0,1),$$

where $f(\bar{n}(A))$ scales with an effective participant base $\bar{n}(A)$ (capped in implementation to preserve numerical stability).

Token supply Token supply evolves through Poisson issuance:

$$X_{t+\Delta t} = X_t + x_D \cdot \mathbf{1}\{\text{issuance event in } (t, t+\Delta t]\},$$

where issuance events arrive with intensity λ_{token}.

Delegate effort choice

Effort costs are specified as quadratic in the leading delegate's effort. In particular, the cost of effort is given by

$$c(e_{\text{lead}}) = \tfrac{1}{2}\, c_2\, e_{\text{lead}}^2,$$

where $c_2 > 0$ governs the marginal cost of effort. The corresponding marginal cost of effort, used in the leading delegate's effort first-order condition, is

$$c'(e_{\text{lead}}) = c_2\, e_{\text{lead}}.$$

This quadratic specification ensures convex effort costs and delivers an interior solution for the delegate's optimal effort choice.

Honest leading delegate The honest leader's effort choice solves

$$c'(e^*_{\text{lead}}) = \mu'_A(e^*_{\text{lead}})\, \frac{\partial W_H(S_t)}{\partial A}.$$

The derivative $\partial W_H(S_t)/\partial A$ is computed locally using a state-dependent approximation, and the FOC is solved numerically at each time step.

Malicious leading delegate and counterfactual effort. When the leading delegate is malicious, effort is determined using the marginal continuation value $\partial W_M(S_t)/\partial A$, which incorporates both the honest continuation value and the option value of optimally attacking in the future. In the implementation, this marginal value is decomposed as

$$\frac{\partial W_M(S_t)}{\partial A} = \frac{\partial W_H(S_t)}{\partial A} + \frac{\partial V_{\text{opt}}(S_t)}{\partial A},$$

where $S_t = (A_t, T_t, X_t, p_{i,t}, p_{j,t})$ denotes the state.

The first term, $\partial W_H(S_t)/\partial A$, is the marginal continuation value of an honest incumbent and is computed using the honest-value derivative routine. The second term captures how productivity affects the expected value of a future attack through its impact on treasury growth. In the model, productivity shifts the treasury drift via the effective participant base $\bar{n}(A)$. Applying the chain rule, the marginal effect of productivity on treasury drift is

$$\frac{\partial \mu_T}{\partial A} = \frac{\partial f(\bar{n}(A))}{\partial \bar{n}} \cdot \frac{\partial \bar{n}(A)}{\partial A},$$

where $f(\bar{n})$ denotes treasury inflows. In the implementation,

$$\frac{\partial \bar{n}(A)}{\partial A} = \frac{a_0\,\varepsilon}{\delta_n}\,A^{\varepsilon-1}, \qquad \frac{\partial f(\bar{n})}{\partial \bar{n}} = f_0\,\beta_f\,\bar{n}(A)^{\beta_f-1},$$

with derivatives evaluated at $A > 0$ and set to zero otherwise.

The option-value component is discounted by the expected time until an attack occurs. Attacks arrive with endogenous hazard

$$h(S_t) = \lambda_{\text{gov}}\, q^*(S_t, n_I^*),$$

where $q^*(S_t, n_I^*)$ is the equilibrium attack probability evaluated at the equilibrium level of information

acquisition n_I^*. The expected time to attack is approximated as

$$\mathbb{E}[\tau \mid S_t] \approx \frac{1}{\lambda_{\text{gov}}\, q^*(S_t, n_I^*)}.$$

Combining these elements, the marginal option value is approximated as

$$\frac{\partial V_{\text{opt}}(S_t)}{\partial A} \approx e^{-r\,\mathbb{E}[\tau]} \cdot \frac{\partial \mu_T}{\partial A} \cdot \mathbb{E}[\tau],$$

which captures the discounted marginal increase in expected treasury accumulation accruing to the malicious delegate prior to attack.

For comparison exercises, we additionally compute counterfactual honest effort along malicious state paths by holding the realized state variables fixed and re-solving the honest effort first-order condition at each date.

Voting subgame and information acquisition

At each date, tokenholders may acquire information and vote in a governance contest. Voting is implemented using Poisson vote arrivals, implying that the net vote margin follows a Skellam distribution. Pivotal probabilities and expected payoff differences imply a marginal benefit of acquiring information; equilibrium information acquisition $n_I^*(S_t)$ solves the marginal-benefit-equals-cost condition subject to the bounds $n_I^*(S_t) \in [0, \bar{n}(A_t)]$. The implied individual information-acquisition probability is

$$\pi_t \equiv \frac{n_I^*(S_t)}{\bar{n}(A_t)}.$$

Governance events, attacks, and belief updating

Governance opportunities arrive as a Poisson process with intensity λ_{gov}. Conditional on a governance event, a malicious incumbent may attack with probability $q_i^*(S_t)$, a reduced-form function of the current

state and the voting equilibrium.

An attack yields payoff

$$V_{\text{attack}}(S_t) = T_t + W_{\text{out,reset}} - c_I.$$

An attack occurs only if this payoff exceeds the incumbent's continuation value and a Bernoulli draw with success probability $q_i^*(S_t)$ realizes. The attack probability for a malicious delegate is specified in reduced form as a logistic function,

$$q_i^*(S_t) = \frac{1}{1+\exp(-\Xi(S_t))},$$

where

$$\Xi(S_t) = 2\log(1+T_t) \;+\; 2\,(1-p_{i,t}) \;-\; \frac{n_I^*(S_t)}{\bar{n}(A_t)}.$$

The first term captures the incentive effect of a larger treasury, which increases the gains from attacking. The second term reflects reputational discipline: lower perceived honesty ($p_{i,t}$) raises attack incentives. The final term captures oversight, as a higher share of informed tokenholders $n_I^*(S_t)/\bar{n}(A_t)$ reduces the probability of attack. The logistic specification ensures $q_i^*(S_t) \in (0,1)$ for all admissible states. When an attack occurs, the treasury experiences a discrete downward jump of size $\Delta T = -T_t$, corresponding to a full extraction of treasury funds, and beliefs are reset according to the model's transition rules.

At each governance opportunity, if no attack is observed, tokenholders update their beliefs about the delegate's type using Bayes' rule. Let p_t^- denote the prior belief that the incumbent delegate is honest, and let $q^*(S_t)$ denote the equilibrium probability that a *malicious* delegate attacks in state S_t.

Conditional on observing no attack, posterior beliefs are updated according to

$$p_t^+ = \frac{p_t^-}{p_t^- + (1-p_t^-)(1-q^*(S_t))}.$$

Simulation algorithm

The full simulation proceeds as follows.

1. Initialize $S_0 = (A_0, T_0, X_0, p_{i,0}, p_{j,0})$.
2. For each grid time t_k:
   (a) Clear the token market to obtain $x_u(S_{t_k})$.
   (b) Solve the voting subgame to obtain $n_I^*(S_{t_k})$ and π_{t_k}.
   (c) Determine leader effort e_{lead,t_k} (or its malicious analogue).
   (d) Update (A, T, X) using the realization of A_t and X_t.
   (e) Check for governance events and possible attacks; update beliefs using Bayesian updating.
3. Record all state variables, equilibrium objects, attack indicators, and event indices.

Tokenholder utility and pricing

The price at state S_t is defined as the marginal value of tokens held by outside investors evaluated at the market-clearing outside quantity $x_u(S_t)$:

$$P(S_t) \equiv \left.\frac{\partial U(S_t, x)}{\partial x}\right|_{x=x_u(S_t)}.$$

Given a simulated path $\{S_{t_k}\}_{k=0}^K$, tokenholder utility is computed as a discrete-time approximation to discounted payoffs until a stochastic liquidity/exit event. Let τ denote the (random) liquidity time, and

let the per-event information cost be c_A. The implementation computes utility as

$$U(S_{t_k}, x) \approx \mathbb{E}_{t_k}\left[e^{-r(\tau - t_k)}\, P(S_\tau)\, x \;-\; \sum_{\ell : t_\ell < \tau} e^{-r(t_\ell - t_k)}\, c_A \cdot \mathbf{1}\{a^*(S_{t_\ell}) = I\}\right],$$

where $a^*(S_{t_\ell})$ denotes the information-acquisition decision implied by the policy used (in the baseline, the equilibrium policy π_ℓ). In computation, τ is mapped to a discrete exit index using the appropriate hazard specification.

Liquidity-hazard integral approximation

A first pricing function interprets the token as a claim that pays the treasury-to-supply ratio upon a Poisson liquidity event with intensity δ. Under continuous time, a standard identity yields

$$P^{\mathrm{LH}}(S_t) \approx \int_0^{s_{\max}} e^{-(r+\delta)s}\, \delta\, \frac{\widehat{T}_{t+s}}{\widehat{X}_{t+s}}\, ds,$$

where $(\widehat{T}_{t+s}, \widehat{X}_{t+s})$ are constructed using a drift-only forecast path. In implementation, $P^{\mathrm{LH}}(S_t)$ is evaluated numerically on a fine grid with step ds up to truncation $s_{\max}$, which controls approximation error from truncating the infinite integral.

Discrete-time dynamic-programming recursion

A second pricing function computes the same liquidity-hazard valuation using a discrete-time dynamic-programming recursion. This approach can be interpreted as a fixed-point approximation to the continuous-time pricing equation and is often numerically more stable than repeated numerical quadrature. Let k index simulation-grid time points and define the per-step probability of a liquidity event as approximately $\delta\,\Delta t$. The token price satisfies the discrete-time Bellman equation

$$P_k^{\mathrm{DP}} = e^{-r\Delta t}\Big[(1 - \delta\Delta t)\, P_{k+1}^{\mathrm{DP}} \;+\; \delta\Delta t \cdot \left(\frac{T_k}{X_k}\right)\Big],$$

which characterizes P_k^{DP} as the fixed point of a one-step pricing operator that maps continuation values into current prices. The recursion is solved backward along the simulation grid, imposing a terminal condition at the final time point (e.g., $P_K^{\mathrm{DP}} = T_K/X_K$, or an equivalent one-step continuation approximation). Iterating yields a price path that is internally consistent with the realized state variables $\{(T_k, X_k)\}_{k=0}^{K}$ and approximates the continuous-time liquidity-hazard pricing formula. As $\Delta t \to 0$, the fixed-point recursion converges to the corresponding continuous-time valuation equation.

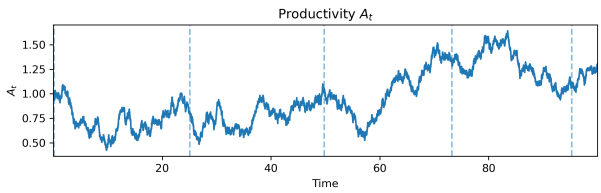
Model parameters

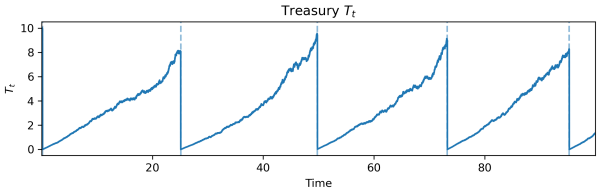
We calibrate key model parameters using empirically plausible orders of magnitude drawn from observed DAO governance activity and market characteristics. First, we calibrate the governance arrival rate λ_{gov} to 1.5 to generate approximately 18 governance proposals per year (if one period is considered to be a month), closely matching the average of 17.6 proposals observed in our sample. This choice ensures that the frequency of governance opportunities in the model aligns with the level of proposal activity faced by tokenholders in practice. Second, delegate compensation is calibrated to reflect the fact that, in most DAOs, remuneration is overwhelmingly paid in governance tokens rather than in dollar-denominated assets. Accordingly, we set token-based compensation to be approximately 100 times larger than dollar-denominated compensation. Third, we calibrate volatility parameters to reflect differences in asset composition. Governance token volatility is set to be roughly twice that of the treasury, consistent with the observation that DAO treasuries hold a mix of relatively stable assets, such as stablecoins, alongside more volatile assets, including Ether and Bitcoin. This asymmetry ensures that attacks and governance events disproportionately affect token prices relative to treasury values. We discipline our calibration using observed growth dynamics in the DeFi sector over the past two years. Aggregate DeFi total value locked (TVL) grew from roughly \$55 billion to \$125 billion, corresponding to an average annual growth rate of approximately 51%. Over the same period, Aave experienced average annual growth of roughly 182% in TVL. Netting out aggregate DeFi market growth implies an idiosyncratic annual growth rate of approximately 87 percent, or 5.35% per month, using the standard transformation

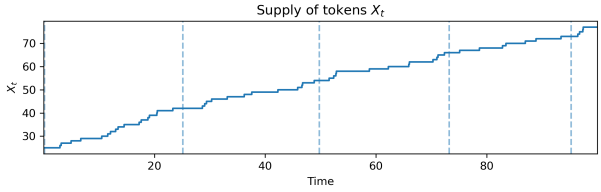
$g_m = (1 + g_a)^{1/12} - 1$. Similarly, Uniswap exhibited annual growth of approximately 156% , after removing the 51% market component, this corresponds to an idiosyncratic annual growth rate of roughly 70%, or 4.5% per month. Thus, across the two most prominent DeFi DAOs, net-of-market monthly growth rates consistently fall in the 4.5–5.5 percent range. Finally, we set the baseline productivity drift to zero, so that productivity growth in the simulations should be interpreted net of aggregate market growth. Under this normalization, all productivity dynamics arise endogenously from delegate effort and governance incentives rather than from exogenous trends in the broader crypto or DeFi market. Motivated by the empirical magnitudes above, we calibrate the idiosyncratic monthly growth rate to 5 percent, which closely matches the observed net-of-market growth of Aave and Uniswap and provides a conservative benchmark for protocol-specific productivity dynamics.

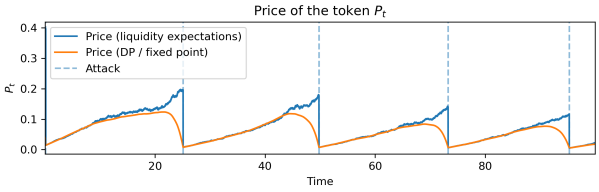
Parameter	Description	Value
r	Discount rate	0.05
δ	Depreciation rate	0.1
λ_{gov}	Governance arrival rate	1.5
$\mu_{0,A}$	Baseline productivity drift	0.0
$\mu_{1,A}$	Enhanced productivity drift	0.05
α_A	Productivity persistence	0.7
σ_A	Productivity volatility	0.1
r_T	Treasury return rate	0.05
f_0	Baseline utility	1.0
β_f	Concavity of the treasury inflow	0.5
ω	Compensation of the delegate	0.1
σ_T	Treasury volatility	0.05
a_0	Attack intensity base	0.05
ϵ	Elasticity parameter	0.5
δ_n^{const}	Constant exit rate	0.5
c_A	Cost of productivity effort	0.05
x_D	Delegate's stake	1.0
λ_{token}	Token arrival rate	0.5
c_2	Quadratic cost parameter	1.0
c_I	Cost of information acquisition	0.1
$p_{0,\text{rep}}$	Initial reputation	0.2
ΔU_{vote}	Utility gain from voting	5.0
$e_{\min}$	Minimum effort	0.0
$e_{\max}$	Maximum effort	1.0

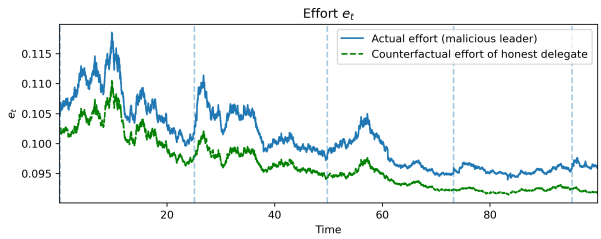
Results of the Model simulation (T = 100 periods)

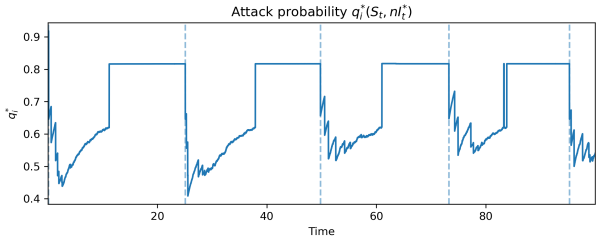


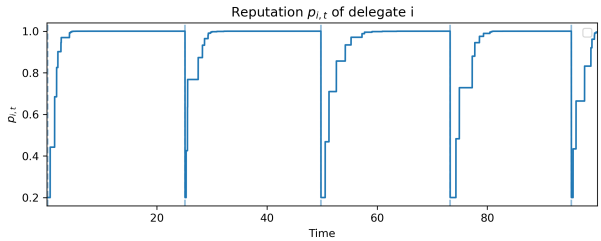


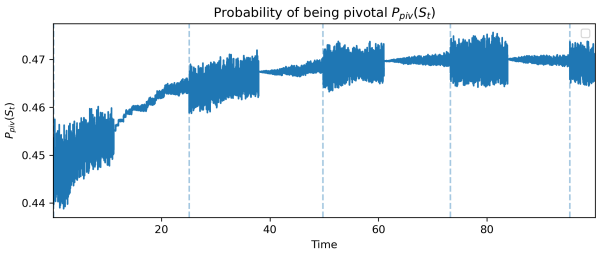


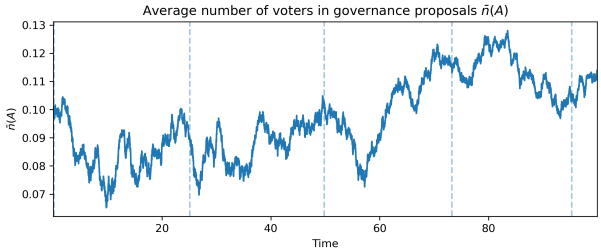


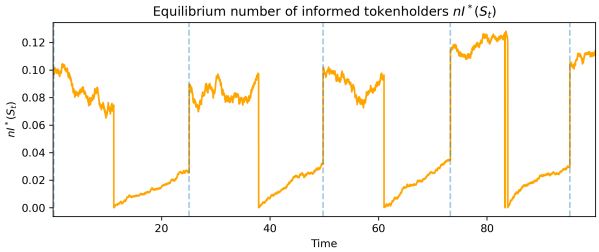




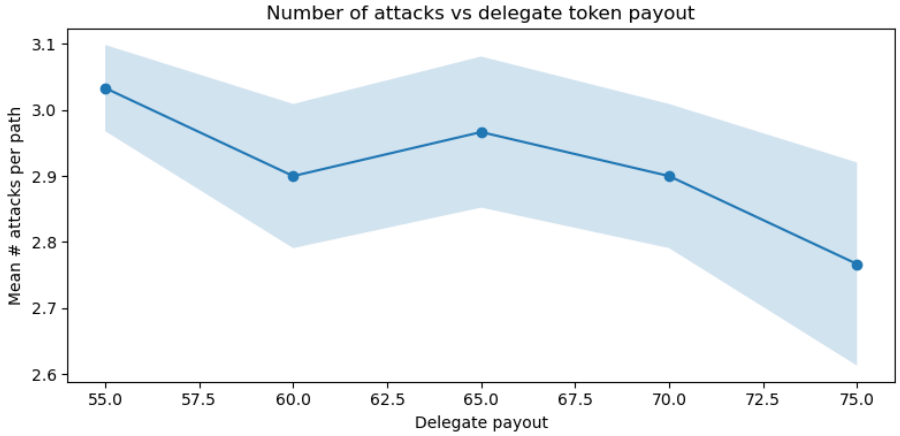








Decreasing number of attacks as the delegate token payout increase (T = 50 periods)



Robustness check: Average and 95% confidence intervals over 50 simulation (T = 100 periods)