The Entry Trilemma: Costly Registration as Sybil-Resistance in Graded-Reward Networks

Withheld for anonymous review

Abstract

Permissionless reward networks must let anyone register an identity, yet permissionless entry invites the Sybil attack: a single participant registering many identities to capture a larger share of reward. We study the canonical defense, a registration cost, in networks that divide a fixed reward across participants in proportion to a strictly increasing function of measured performance. This score-proportional class includes the softmax-over-ratings rule used to allocate emissions on decentralized machine-intelligence subnets. Our main result is an impossibility theorem, the *Entry Trilemma*: no mechanism in this class can simultaneously provide free entry, Sybil-resistance, and score-proportional grading; any two preclude the third. The boundary is sharp: the impossibility holds because reward tracks measured performance, which a copy reproduces at no cost, and it dissolves only by leaving that class, which abandons fixed-pool performance-based reward. The proof is constructive and exact. Under free entry, cloning the leading strategy is strictly profitable in every score-proportional rule. A registration burn restores Sybil-resistance without abandoning score-proportional reward, and the smallest burn that deters cloning equals the dilution a copy imposes on incumbents, which makes it a Pigouvian congestion price rather than a tax. Charging that price in full has a cost we make precise: because reward shares sum to one, the burn that eliminates cloning prices out all but a handful of participants, so clone-resistance and broad participation cannot be maximized together, a quantitative counterpart of the impossibility. A demand-responsive burn that bounds cloning while keeping the board populated is the interior compromise the trade-off forces. Because the clone-resistant level moves with demand and is never observed, we prove the deployed rule, which raises the burn on each entry and lets it decay otherwise, stays within a fixed factor of that moving target in the worst case, a competitive guarantee no constant burn meets. A *Toll-Reduction Paradox* follows: any populated board sits below the clone-resistant level, so lowering the burn cannot reduce cloning and only deepens the duplication of the highest-value strategies. The registration cost is thus a position on a frontier between resisting duplication and admitting entry, not a free parameter a too-high-cost objection can move to everyone's benefit. Evidence from all 128 live subnets of one such network is consistent with the theory: the most-demanded subnet carries a registration price many times the next-highest and tens of thousands of times the median, while 82 percent of subnets sit at or below the threshold commonly called free, precisely because they attract no entry. On this evidence the most-demanded subnet already sits below the level that would deter cloning, so the case for lowering its toll runs backwards.

Keywords: Sybil resistance, registration cost, contest design, softmax, congestion pricing, mechanism design, token networks, *JEL:* C72, D44, D62, D82, H23

1. Introduction

A decentralized reward network pays participants for measured contributions. To remain permissionless it must let anyone register a working identity. The same openness that makes the network credible also exposes its central vulnerability: the Sybil attack, in which one participant registers many identities to harvest a larger share of a fixed reward pool (Douceur, 2002). The standard defense is to make registration costly, by proof of work, a stake, or a token burn. The defense is ubiquitous, but its economics are rarely stated precisely. Why is any positive entry cost *necessary* rather than a deadweight tax on participation? How large should it be? And when participants object that the cost is too high, is the objection a correctable design flaw or a structural feature?

We answer these questions for the class of mechanisms that allocate a fixed reward in proportion to a strictly increasing function of each participant's measured performance. We call this the *score-proportional* class. It is the relevant class in practice: the emission rule on the subnets we study divides each epoch's reward by a softmax over performance ratings, and proportional sharing is the natural functional form whenever the designer wants reward to track demonstrated quality smoothly rather than award a single winner.

Our central result is an impossibility theorem. Three properties that a designer might want, free entry, Sybil-resistance, and score-proportional grading, cannot hold together. Any two are achievable; the third must be sacrificed. We call this the Entry Trilemma (Theorem 1). The proof is not an existence argument: we exhibit the profitable deviation explicitly. Under free entry, an agent who holds the leading strategy strictly increases total reward by registering a behavioral copy of it, in *every* score-proportional rule (Proposition 1). The deviation cannot be deterred by detecting and removing copies, because behavioral copies survive paraphrase; it is deterred only by pricing entry. The boundary is sharp, and it is the substantive content of the restriction to score-proportional rules: the impossibility holds because reward tracks measured performance, which a copy reproduces at no incremental cost, so cloning multiplies an agent's share. It dissolves only by leaving that class, either by tying reward to a rival per-identity resource, a bid or a stake, that copies must divide, or by shrinking the total reward as identities multiply; both escapes, used in recent constructions we discuss in Section 2, abandon fixed-pool performance-based reward and are unavailable to a network whose purpose is to pay for demonstrated quality out of a fixed emission.

Pricing entry is exactly what a registration burn does. We show that the smallest burn that removes the cloning incentive equals the dilution a copy imposes on incumbents, which makes the burn a Pigouvian congestion price (Pigou, 1920; Vickrey, 1969) rather than a tax (Proposition 2). Charging it in full, however, prices out every seat worth less than the clone's gain, and because reward shares sum to one only a handful of seats clear that bar: clone-resistance and broad participation cannot both be maximized (Proposition 3), a quantitative echo of the trilemma. The cost that keeps the board populated must therefore sit below the level that eliminates cloning, bounding it instead; and because entry demand is unknown and arrives in bursts, that margin moves over time, so no constant burn stays on it (Proposition 5). The adaptive burn that rises with each registration and decays when entry pauses is its natural implementation, and we prove it stays within a fixed factor of the unobserved, moving optimum in the worst case while no constant burn does, a competitive-analysis counterpart to the static impossibility (Theorem 2).

We then study the comparative statics of the toll, motivated by a common objection that the registration cost is too high and should be lowered. We prove a Toll-Reduction Paradox (Proposition 7): from any populated board, where the burn already sits below the clone-resistant level, lowering it cannot reduce cloning. A copy of a strategy of share w becomes profitable once the burn falls below $Vw/(1+w)$, a threshold largest for the highest-share strategies, so a cut deepens duplication at the top of the board, where the network's value is concentrated, and each copy consumes a seat while adding no new strategy. A corollary frames the governance implication (Corollary 2): the registration cost sits on a frontier between resisting duplication and admitting entry, so the too-high-cost objection cannot be met by lowering the cost without reviving the duplication the cost exists to suppress. The adjustments that help a dissatisfied participant, raising performance or entering a less contested venue, lie outside the cost parameter.

Finally we bring the theory to data. Using a snapshot of all 128 live subnets of one such network, we document that registration cost is extraordinarily dispersed: the single most-demanded subnet prices entry at many times the next-highest subnet and tens of thousands of times the median, while 82 percent of subnets sit at or below the threshold commonly described as free. The cheap subnets are cheap because almost no one is trying to enter them: without registrations the burn decays to its floor, whatever a subnet's task or parameters. The most-demanded subnet sits orders of magnitude above that floor because entry there is both persistent and priced on the aggressive adaptive schedule such demand warrants. We also show, on the most-demanded subnet, that the top strategies by rating fall into a single near-identical cluster under a text-similarity measure, even though every one of them paid the highest entry price in the network; at zero price the convergence would be complete. Because the strategies are public, this convergence is the predicted competitive response of the whole field to a readable leader, not the act of any single participant. Because a populated board already sits below the clone-resistant level, this convergence is what a sub-threshold toll necessarily leaves in place; lowering the toll could only deepen it.

The paper proceeds as follows. Section 2 places the results in the literature. Section 3 sets up the model. Section 4 proves the Entry Trilemma. Section 5 derives the optimal burn and its adaptive form. Section 6 explains why the reward gradient is interior rather than winner-take-all. Section 7 proves the Toll-Reduction Paradox and its governance corollary. Section 8 presents the cross-network evidence. Section 9 explains why a within-market experiment cannot settle the dispute and points to the companion paper. Section 10 concludes.

2. Related work

Sybil resistance and entry costs. The Sybil attack was named by Douceur (2002), who showed that without a trusted identity authority a logically centralized entity can present arbitrarily many identities, and that resource-bound entry tests are the only general defense in an open system. Permissionless networks operationalize the resource bound through proof of work, bonded stake (Saleh, 2021), or a token burn (Karantias et al., 2020).

Sybil-proofness against graded reward. That paying more than a single winner is in tension with Sybil-proofness is known across several settings. For direct revelation mechanisms, Pan et al. (2024) show that the only non-wasteful, symmetric, dominant-strategy incentive-compatible and Sybil-proof rule is the second-price auction, a winner-take-all object; they also observe that a weaker Bayesian Sybil-proofness can admit graded rules, which is why our solution concept below is dominant strategies. For reward trees and referral chains, Babaioff et al. (2012), Emek et al. (2011), and Lv and Moscibroda (2013) prove analogous split-proofness impossibilities for multi-recipient payouts. In restaking, Chitra et al. (2025) study how pooled versus winner-take-all rewards shift Sybil incentives, the same gradient-versus-Sybil axis. Platt et al. (2024) state a “Sybil attack vulnerability trilemma” among permissionlessness, Sybil-resistance, and zero participation cost; their third leg is the entry cost itself. Our setting differs on three counts that matter for the result: reward is an emission pool paid *out* for measured performance, not a payment collected in an auction; the relaxing instrument is an explicit, adaptive registration burn whose clone-resistant level we characterize; and the third leg of our trilemma is the shape of the reward gradient, not merely the presence of a cost.

When free entry and grading coexist. The tension is conditional, and naming the condition is part of the contribution. Garimidi et al. (2026) construct a free-entry, Sybil-proof, non-winner-take-all procurement mechanism in which reward is proportional to a per-identity *costly* resource, a bid, so that splitting it across identities is share-neutral. Mazorra and Della Penna (2023) take a different route, a graded multi-recipient schedule that is Sybil-proof at zero entry cost because the total payout *shrinks* as recipients multiply, removing the cloning gain by relaxing the fixed pool rather than by pricing entry. Both routes lie outside the maintained assumptions here, which fix the reward pool and tie reward to measured performance rather than to a rival resource. This is exactly the boundary of Theorem 1: within a fixed pool paid for performance, the impossibility holds because a copy reproduces measured performance at no incremental cost, and it dissolves only by leaving that class, either by tying reward to a rival resource that copies must divide or by shrinking the pool so that copies are self-defeating. A network whose reward must track demonstrated quality cannot take that escape without ceasing to reward quality, so in that setting the burn is the instrument that remains. A close analogue is costly voting: Wagman and Conitzer (2008) show the optimal false-name-proof rule is degenerate at zero cost and that a positive per-identity cost restores responsiveness, a cost-relaxes-degeneracy result in the spirit of ours. Finally, Moulin (2008) shows the proportional rule *is* split-proof when a fixed claim is partitioned; the difference here is that entry is not a partition of a fixed claim but the arrival of a fresh claimant that dilutes the pool, which is what makes cloning profitable.

Contest and tournament theory. Reward that is increasing in relative performance is a contest. The optimal allocation of prizes across ranks is studied by Lazear and Rosen (1981), Moldovanu and Sela (2001), Glazer and Hassin (1988), and Sisak (2009), and the all-pay auction limit by Tullock (1980) and Baye et al. (1996). A central theme, the discouragement effect by which concentrating the prize on the top suppresses the effort of everyone else, is due to Szymanski and Valletti (2005). We use these results to locate the deployed softmax temperature: a graded prize structure is effort-optimal for a heterogeneous field, and both the winner-take-all and the uniform limits are dominated. Our novel contribution to this literature is not the optimal prize schedule but the interaction of the schedule with *entry*: the very gradedness that contest theory recommends is what creates the cloning incentive that the burn must price. Closest to our setting, Garimidi et al. (2025, 2026) bring Tullock-style tolls to blockchain entry to escape winner-take-all, but analyze a toll fixed in advance; Theorem 2 treats the toll as a controller and shows the deployed demand-responsive rule stays within a fixed factor of the moving optimum that no fixed toll attains, a competitive-analysis result (Borodin and El-Yaniv, 1998) new to this literature.

Congestion pricing. The result that the welfare-restoring price of a congesting activity equals its marginal external cost is Pigou (1920), with the dynamic and capacity-investment treatment in Vickrey (1969). We identify the registration burn as a Pigouvian toll on a specific externality: the reward a marginal identity dilutes away from incumbents. The corrective property is that the entrant faces a price equal to the dilution it imposes; because the burn is recycled rather than destroyed, it is a transfer in the Pigouvian sense, and like any Pigouvian toll it need not be rebated to the diluted incumbents to correct the entry margin.

Signaling, cheap talk, and skin in the game. That a costless message is uninformative in equilibrium while a costly one can separate types is the lesson of Crawford and Sobel (1982) and Spence (1973). That decision rights should carry symmetric downside is the agency-theoretic point of Jensen and Meckling (1976) and Holmström (1979). These ideas inform our reading of the objection, and are developed into a governance mechanism in the companion paper (Anonymous, 2026). The behavioral backdrop, that costless voice aggregates into cascades and groupthink, is Banerjee (1992), Bikhchandani et al. (1992), Hirschman (1970), Olson (1965), and Janis (1972).

3. Model

A network runs in discrete epochs. In each epoch a fixed reward, normalized to 1, is divided among occupied *seats*. There are M seats (a capacity). An *agent* controls one or more seats. Each seat runs a *strategy* s drawn from a strategy space $\mathcal{S}$. A strategy induces a measured *performance* (a rating) $p(s) \in \mathbb{R}$; two seats running the same strategy realize the same performance distribution. Let $g : \mathbb{R} \to \mathbb{R}_{>0}$ be a strictly increasing *score function*.

Definition 1 (Score-proportional mechanism)**.** Given occupied seats $1, \ldots, n$ with performances $p_1, \ldots, p_n$, a score-proportional mechanism pays seat i

$$\phi_i \ = \ \frac{g(p_i)}{\sum_{j=1}^{n} g(p_j)}, \qquad \sum_{i=1}^{n} \phi_i = 1.$$

Entry costs $c \geq 0$ per seat, paid once at registration. An agent's payoff is $V \sum_{i \in A} \phi_i - c\,|A|$, where A is the set of seats it controls and $V > 0$ is the present value of one epoch-unit of reward across the seat's lifetime.

The softmax-over-ratings rule used in practice is the special case $g(p) = e^{p/T}$ for a temperature $T > 0$. Linear proportional sharing is $g(p) = p$ on $\mathbb{R}_{>0}$. We make one structural assumption, which captures the open nature of the network.

Assumption 1 (Anonymity / no identity oracle)**.** The mechanism conditions only on the vector of measured performances. It does not observe which agent controls which seat, and it cannot certify that two seats run the same underlying strategy. Equivalently, a behavioral copy of a strategy, a paraphrase that performs identically, is indistinguishable to the mechanism from an independent entry.

Assumption 1 is the realistic condition for a permissionless network. Text-level deduplication is defeated by paraphrase; behavioral fingerprinting is itself gameable; and a trusted identity authority would contradict permissionlessness. We discuss in Section 4 why pricing entry dominates detection precisely because pricing does not require the oracle that Assumption 1 denies.

Assumption 2 (Non-rivalry of performance)**.** A behavioral copy of a strategy attains the same performance as the original at no incremental per-identity cost, and without reducing the original's performance. Performance is reproduced by cloning, not divided.

Assumption 2 separates performance-weighted reward from resource-weighted reward, and it is what makes cloning profitable below. A strategy that debates well debates equally well when copied: skill is non-rival to the copier. A bid or a stake is rival: an agent with a fixed budget who spreads it across copies funds each one less, so a reward share linear in that resource is unchanged by splitting. The score function g in Definition 1 is applied to measured performance, not to a funded resource, which is the substantive restriction of the score-proportional class and the hinge of the boundary discussed after Theorem 1.

We study three properties.

Definition 2 (The three properties)**.** (A) *Free entry:* $c = 0$.

(B) *Sybil-resistance:* for every profile and every agent, registering one additional seat that runs a strategy already present does not strictly increase the agent's payoff.

(C) *Score-proportional grading:* the mechanism is score-proportional in the sense of Definition 1 for a strictly increasing g, so reward strictly increases in own performance and more than one performance level receives positive reward.

Entry and welfare. The score-proportional mechanism is the second stage of a two-stage game; the first stage is entry. A population of potential entrants chooses whether to register against the seat capacity M. An *honest* entrant is endowed with a *distinct* strategy whose performance p is drawn from a distribution F on $[\underline{p}, \overline{p}]$; distinct strategies are the output the network exists to elicit. A *cloner* contributes no new strategy but registers a behavioral copy of the highest-performing strategy already present, which Assumption 2 lets it do at the same performance.

Facing a burn c and the field its rivals form, an honest type p registers iff its expected payoff is nonnegative,

$$V \, \frac{g(p)}{Z} - c \ \geq \ 0, \tag{PC}$$

with $Z = \sum_j g(p_j)$ the equilibrium normalizer. An entry equilibrium is an occupied set in which every occupant satisfies (PC) and no excluded honest type strictly prefers to enter; when capacity binds, the seats are taken by the highest-performing types that satisfy (PC). The marginal admitted type p_c solves $Vg(p_c)/Z = c$, so a higher burn raises p_c and weakly shrinks the honest field. Two regimes recur below: *slack* capacity, in which the burn sets the entry margin, and *binding* capacity with honest demand in excess of the M seats, in which the burn rations a queue. The most-demanded subnet of Section 8 is in the second regime: its seats are full and entry demand persists.

Definition 3 (Network value)**.** Network value is the total social value of the *distinct* strategies in play,

$$W = \sum_{s \in D} v\big(p(s)\big),$$

where D is the set of distinct strategies occupying seats and v is strictly increasing. A behavioral copy occupies a seat but adds no element to D, so it contributes nothing to W. The registration burn, which on the network we study is recycled into future emission rather than destroyed, is a transfer and does not enter W. For distributional statements we track each honest participant's reward share $w_i = g(p_i)/Z$ directly rather than an aggregate net surplus, since the transfer's incidence depends on registration timing and so cancels in W but not in any one private payoff.

The single modeling commitment in Definition 3 is that the network is paid for, and produces, *distinct* contributions: a redundant copy of an existing strategy produces nothing the original did not. This is the non-rivalry of Assumption 2 read on the output side.

4. The Entry Trilemma

Theorem 1 (Entry Trilemma)**.** *Under Assumptions 1 and 2, with Sybil-resistance (B) required in dominant strategies, no mechanism satisfies (A), (B), and (C) simultaneously. Each pair is achievable, so exactly one property must be relinquished.*

Proof. We show (A) and (C) imply the negation of (B). Fix any profile with at least two occupied seats whose performances are not all equal, and let $p^* = \max_i p_i$ be the top performance, attained by a seat controlled by some agent. Because $g > 0$, the normalizer $Z = \sum_j g(p_j)$ satisfies $Z > g(p^*)$ whenever any other seat is occupied. The agent's reward from its top seat is $g(p^*)/Z$.

Under (A) the agent registers, at zero cost, one more seat running a behavioral copy of the top strategy. By Assumption 1 the copy is treated as an independent entry with performance p^*, so the new normalizer is $Z' = Z + g(p^*)$ and the agent now controls two seats each paid $g(p^*)/Z'$. The change in the agent's total reward is

$$\Delta \;=\; \frac{2\,g(p^*)}{Z + g(p^*)} \;-\; \frac{g(p^*)}{Z} \;=\; \frac{g(p^*)\,\big[\,Z - g(p^*)\,\big]}{Z\,\big[\,Z + g(p^*)\,\big]} \;>\; 0,$$

because $Z - g(p^*) > 0$. With $V > 0$ and $c = 0$ the agent's payoff strictly increases, so (B) fails.

Each pair is achievable. Dropping (A): a sufficiently large registration burn makes the deviation above unprofitable while the mechanism remains score-proportional, so (B) and (C) hold; Proposition 2 gives the exact threshold (Section 5). Dropping (B): a network that simply tolerates Sybils satisfies (A) and (C). Dropping (C): a winner-take-all rule, which pays the entire reward to a single top seat, satisfies (A) and (B), since once an agent holds a top seat additional copies add nothing (the prize is already captured). Hence all three cannot hold, but any two can. ■

Remark 1 (The resource-proportional escape, and why it is closed here). Assumption 2 is essential, not cosmetic. If reward were proportional to a per-identity costly resource, a bid or a stake, rather than to measured performance, a cloning agent would have to split that resource across its copies, and under a share linear in the resource the split would be exactly neutral. Free entry, Sybil-resistance, and a graded multi-recipient schedule can then coexist, as Garimidi et al. (2026) and Mazorra and Della Penna (2023) establish by construction. That escape is closed for a network whose reward must track demonstrated quality: skill is non-rival, so a copy reproduces the top performance for free and the controlled share multiplies rather than splits. Rewarding performance is what forces the choice in Theorem 1. The result is stated for Sybil-resistance in dominant strategies; under a Bayesian solution concept a graded rule can satisfy a weaker Sybil-proofness in expectation (Pan et al., 2024), which is not a guarantee a permissionless network can rely on against a deliberate attacker.

Remark 2 (Why detection does not rescue free entry). One might hope to keep (A) and (C) by detecting and discarding copies. Assumption 1 is exactly the statement that this fails in an open network: a paraphrased strategy performs identically yet is textually distinct, and no deployable oracle separates a genuine independent entrant who arrived at the same strategy from a deliberate copy. Pricing dominates detection because a burn deters the copy regardless of whether it could ever be identified. The burn is detection-free Sybil-resistance.

Remark 3 (The prize-splitting corner is a sacrifice of (C)). A rule that fixes a reward mass per performance *rank* and splits it equally among seats tied at that rank does deter cloning the top, since copies split rather than multiply the top mass. But such a rule is not score-proportional: a seat's pay depends on how many others share its rank, not only on a common normalizer, and genuine improvements within a rank earn nothing. It is a rank-prize mechanism, a move along the (C) axis toward winner-take-all, and it reintroduces the discouragement effect of Section 6. It escapes the trilemma only by giving up the smooth grading that (C) names. This is consistent with, not a counterexample to, Theorem 1.

5. The optimal burn and why it is adaptive

The proof of Theorem 1 already contains the deployed instantiation.

Proposition 1 (Cloning under the softmax). *For the softmax rule $g(p) = e^{p/T}$, write $w^* = e^{p^*/T}/Z$ for the top-strategy share. Cloning the top strategy at zero cost is strictly profitable, and the gain depends on whether the cloner already holds a seat. A* new identity *that copies the top strategy earns the gross share*

$$\gamma \;=\; \frac{w^*}{1+w^*},$$

while an agent who already holds the top seat *and registers one more copy gains only the smaller net amount*

$$\Delta \;=\; \frac{w^*\,(1-w^*)}{1+w^*} \;=\; (1-w^*)\,\gamma,$$

because the second seat dilutes the first. Both are strictly positive whenever any other seat carries positive weight, and $\gamma > \Delta$ *always, so the most profitable clone, and hence the binding constraint on any clone-resistant cost, is a fresh copy of the leader. Numerically, at the representative top share* $w^* = 0.294$ *implied by the deployed temperature (Section 6),* $\gamma \approx 0.227$ *and* $\Delta \approx 0.16$*, and successive copies push the controlled share toward* 1*.*

Proof. A new identity running the top performance adds $g(p^*) = w^*Z$ to the normalizer, giving $Z' = Z(1 + w^*)$, and earns $g(p^*)/Z' = w^*/(1 + w^*) = \gamma$ from a base of zero. An agent already holding the top seat instead moves from one seat earning $g(p^*)/Z$ to two earning $2g(p^*)/Z'$, a net change $\Delta = g(p^*)[Z - g(p^*)]/(Z[Z + g(p^*)]) = w^*(1 - w^*)/(1 + w^*) = (1 - w^*)\gamma < \gamma$. At $w^* = 0.294$: $\gamma = 0.294/1.294 \approx 0.227$ and $\Delta = 0.294 \cdot 0.706/1.294 \approx 0.16$. ■

Sybil-resistance requires that no copy pay for itself. By Proposition 1 the most profitable copy is a fresh identity duplicating the leader, earning the gross share γ, so the binding condition is that the burn deter it.

Proposition 2 (The clone-resistant burn is a congestion price)**.** *Let* $\gamma = w^*/(1 + w^*)$ *be the worst-case clone gain of Proposition 1 and let* V *be the lifetime value of the entire reward stream, so a seat of share* w *is worth* Vw*. Then (i) a registration burn deters every clone of the top strategy if and only if* $c \geq V\gamma$*; and (ii) the threshold* $V\gamma$ *equals the dilution that a fresh copy imposes on the incumbents as a group. The burn is therefore a congestion price: at* $c^* = V\gamma$ *the marginal identity pays exactly the externality its entry creates, a Pigouvian toll rather than a tax.*

Proof. (i) By Proposition 1 the most profitable clone is a fresh copy earning gross $V\gamma$; it enters iff $V\gamma - c > 0$, so $c \geq V\gamma$ deters it and, since $\gamma > \Delta$, every less profitable clone as well. (ii) A fresh copy of the top raises the normalizer from Z to $Z(1 + w^*)$, so the incumbents' collective reward share falls from 1 to $1/(1 + w^*)$, a loss of $w^*/(1 + w^*) = \gamma$. By conservation of the fixed reward this loss is exactly the copy's own share γ: the entrant's private gain equals the dilution it imposes on the rest of the field. Setting $c^* = V\gamma$ charges the entrant that externality, the Pigouvian property. ■

Charging the externality in full carries a cost the toll cannot escape. Because reward shares sum to one, only a few seats can ever be valuable enough to clear the clone-resistant price.

Proposition 3 (Clone-resistance rations participation)**.** *At most $\lfloor 1/\gamma \rfloor$ seats can carry a reward share of at least γ. A seat is worth holding under burn c only if its lifetime value Vw_i covers the burn, so at the clone-resistant level $c = V\gamma$ every seat with share below γ is priced out and at most $\lfloor 1/\gamma \rfloor$ remain. Full clone-resistance and broad participation therefore cannot hold together: the burn that eliminates top-cloning collapses the board to its few highest-valued seats. Conversely, any burn under which more than $\lfloor 1/\gamma \rfloor$ seats stay occupied is strictly below $V\gamma$ and so bounds cloning without eliminating it.*

Proof. The shares $w_i = g(p_i)/Z$ are nonnegative and sum to one, so at most $\lfloor 1/\gamma \rfloor$ of them can be at least γ. A seat participates under burn c iff its lifetime reward covers the one-time burn, $Vw_i \geq c$; at $c = V\gamma$ this requires $w_i \geq \gamma$, leaving at most $\lfloor 1/\gamma \rfloor$ seats. At the representative share $w^* = 0.294$, $\gamma \approx 0.227$, so the bound is $\lfloor 1/\gamma \rfloor = 4$. ■

Proposition 3 is the Entry Trilemma in quantitative form. The registration cost that buys complete Sybil-resistance does so only by all but eliminating entry, the property (A) the trilemma says must be surrendered. A network that wants both a populated board and resistance to cloning cannot reach either corner; it must sit at an interior burn that bounds cloning rather than removing it. Section 8 shows the deployed burn does exactly this: it keeps all 64 seats occupied. Since $\lfloor 1/\gamma \rfloor$ is a handful for any concentrated field, sixty-four filled seats lie well below $V\gamma$ whatever the precise top share, and in a field concentrated enough that only the leader carried a share above γ the clone-resistant burn would leave a single occupied seat of the sixty-four, the rationing tension at its sharpest. The residual convergence among the top strategies is the cloning that a sub-$V\gamma$ toll necessarily leaves in place.

The multiplicative form of the deployed burn adds a further defense against a *sustained* cloning attack, making it expensive at a rate no constant price can match.

Proposition 4 (Exponential capture cost)**.** *Consider an attacker who registers k clone seats of the leading strategy in immediate succession. Under a burn that starts at b_0, multiplies by a factor $m > 1$ on each registration, and is capped at $c_{\max}$, the total cost of the burst is*

$$C_{\text{mult}}(k) = \sum_{t=0}^{k-1} \min\bigl(b_0\, m^t,\, c_{\max}\bigr),$$

which equals $b_0\,(m^k - 1)/(m-1)$ while the cap is slack, growing geometrically in k, and adds $c_{\max}$ per seat once the cap binds. Under any burn bounded above by a constant $\bar{c} < c_{\max}$, the same burst costs at most $\bar{c}\,k$, linear in k. The adaptive burn is therefore more expensive to the attacker by a factor that *grows without bound in k before the cap and is at least $c_{\max}/\bar{c}$ after it.*

Proof. The burst cost is the sum of the per-registration burns. With multiplicative growth the unclamped terms are geometric, $\sum_{t=0}^{k-1} b_0 m^t = b_0(m^k - 1)/(m-1)$; once the burn reaches $c_{\max}$ each further seat costs $c_{\max}$. A burn bounded by $\bar{c}$ contributes at most $\bar{c}$ per registration, so k seats cost at most $\bar{c}k$. Before the cap the ratio of the geometric cost to $\bar{c}k$ diverges in k; after the cap each seat costs $c_{\max}$, so the per-seat ratio is at least $c_{\max}/\bar{c}$. ■

Section 8 quantifies this gap for the deployed parameters: capturing all 64 seats costs roughly 6,200 TAO under the live doubling burn against 6.4 TAO under a flat 0.1 TAO price (Figure 5).

Proposition 5 (No constant burn tracks the clone-resistant level)**.** *Let the value V_t and the field vary across epochs, so the clone-resistant level $c_t^* = V_t\gamma_t$ is state-dependent and not known in advance. If c_t^* is non-constant, no constant burn $\bar{c}$ equals it in every state: where $c_t^* > \bar{c}$ the top strategy is profitably cloned, and where $c_t^* < \bar{c}$ the burn over-charges entry, rationing participation more tightly than clone-resistance requires. The burn that stays at the clone-resistant margin must therefore move with realized demand.*

Proof. Fix a constant burn $\bar{c}$. Since c_t^* is non-constant, either $\bar{c} < \max_t c_t^*$, and in a state with $c_t^* > \bar{c}$ a fresh copy of the top is profitable by Proposition 2; or $\bar{c} \geq \max_t c_t^*$, and then since c_t^* varies there is a state with $c_t^* < \bar{c}$, where by Proposition 3 the burn prices out every seat with share in $[\gamma_t, \bar{c}/V_t)$, beyond those clone-resistance requires. Only a state-contingent burn equal to c_t^* avoids both, and it is feasible because it is defined pointwise. ■

Because demand is revealed only through registrations, an implementable toll must respond to that signal alone. A burn that multiplies by a fixed factor on each registration and decays geometrically toward a floor between registrations is such a rule: each registration is current evidence of entry pressure and raises the toll, while a pause relaxes it over a fixed horizon, the discrete analogue of a congestion price that responds to instantaneous load (Vickrey, 1969). The deployed schedule is one instance, and Proposition 4 shows the multiplicative form has a property no constant toll shares at any feasible level: it makes a sustained capture attack geometrically, rather than linearly, expensive.

The deployed schedule does more than make capture expensive: it tracks the moving clone-resistant toll about as closely as any rule can that sees only registrations. Model entry block by block. One candidate entrant arrives per block $t = 0, 1, 2, \ldots$, the granularity at which the chain admits registrations. The clone-resistant toll $c_t^* = V_t \gamma_t$ drifts with demand and is never observed by the mechanism. A candidate registers exactly when a fresh copy of the leader is profitable, that is when $b_t < c_t^*$ (Proposition 2); a registration multiplies the burn by $m > 1$ and an idle block decays it by $\delta \in (0, 1)$, both clamped to $[\ell, c_{\max}]$,

$$b_{t+1} = \begin{cases} \min(c_{\max},\, m\, b_t), & b_t < c_t^* \quad \text{(entry)}, \\ \max(\ell,\, \delta\, b_t), & b_t \geq c_t^* \quad \text{(idle)}. \end{cases}$$

Measure performance by the mispricing $\rho_t = \max\{b_t/c_t^*,\ c_t^*/b_t\} \geq 1$, which equals one only when the burn sits exactly at the clone-resistant toll; a rule is *κ-competitive* if $\rho_t \leq \kappa$ at every block. Demand is *slowly varying* on a horizon if $c_t^* \leq c_{t+1}^* \leq m\, c_t^*$ there, meaning it does not fall and does not rise by more than a single registration's worth in one block. On the network we study a block is twelve seconds while c^* moves on the scale of days, so the condition is slack by orders of magnitude; a sudden jump is the subject of Corollary 1.

Theorem 2 (Adaptive tracking)**.** *Let the clone-resistant toll lie in* $[c_{\mathrm{lo}}, c_{\mathrm{hi}}] \subseteq [\ell, c_{\max}]$*, and let* $\delta \geq 1/m$*, meaning one idle block removes no more than a registration adds; the deployed doubling burn satisfies this at any half-life of at least one block.*

- (i) *If demand is slowly varying and* $b_0 \leq m\, c_0^*$*, the multiplicative-increase, geometric-decay burn is* m*-competitive:* $\rho_t < m$ *at every block, a bound independent of the demand range* $c_{\mathrm{hi}}/c_{\mathrm{lo}}$*.*
- (ii) *Every constant burn* $\bar{c}$ *satisfies* $\sup_t \rho_t \geq \sqrt{c_{\mathrm{hi}}/c_{\mathrm{lo}}}$*, the minimum attained at* $\bar{c} = \sqrt{c_{\mathrm{lo}} c_{\mathrm{hi}}}$*; no constant is better than* $\sqrt{c_{\mathrm{hi}}/c_{\mathrm{lo}}}$*-competitive.*

Whenever demand varies over a range exceeding m^2*, the adaptive burn strictly dominates every constant burn in worst-case mispricing.*

Proof. Part (i), over-pricing. We show $b_t < m\, c_t^*$ for all t by induction; it holds at $t = 0$ by hypothesis. Suppose $b_t < m\, c_t^*$. On an entry block $b_{t+1} = \min(c_{\max}, m\, b_t) \leq m\, b_t < m\, c_t^* \leq m\, c_{t+1}^*$, the last step because demand does not fall. On an idle block $b_{t+1} = \max(\ell, \delta b_t)$, and both arguments are below $m\, c_{t+1}^*$: $\delta b_t < b_t < m\, c_t^* \leq m\, c_{t+1}^*$, while $\ell \leq c_{t+1}^* < m\, c_{t+1}^*$. Hence $b_t/c_t^* < m$ throughout. *Part (i), under-pricing.* Whenever $b_t < c_t^*$ the block is an entry and the cap is slack, so $b_{t+1} = m\, b_t$: the burn rises by the factor m on every under-priced block and reaches the band $[c_t^*, m\, c_t^*)$ within $\lceil \log_m(c_t^*/b_t) \rceil$ registrations. Because demand rises by at most m per block, once in the band the burn stays in it: an idle block can drop b just below c^*, but a single registration then restores it, and at that block $c_t^*/b_t \le 1/\delta \le m$. Combining the two sides, $\rho_t < m$ once the band is reached, and the argument used no feature of the range. *Part (ii).* A constant $\bar{c}$ meets both ends of the demand range, so $\sup_t \rho_t = \max(\bar{c}/c_{\text{lo}},\, c_{\text{hi}}/\bar{c})$. The larger of two positive numbers is at least their geometric mean, $\sqrt{(\bar{c}/c_{\text{lo}})(c_{\text{hi}}/\bar{c})} = \sqrt{c_{\text{hi}}/c_{\text{lo}}}$, with equality iff $\bar{c}/c_{\text{lo}} = c_{\text{hi}}/\bar{c}$, that is $\bar{c} = \sqrt{c_{\text{lo}} c_{\text{hi}}}$. The final claim compares m with $\sqrt{c_{\text{hi}}/c_{\text{lo}}}$, which exceeds m once $c_{\text{hi}}/c_{\text{lo}} > m^2$. ■

The competitive constant is the increase factor m itself, the multiplier behind the doubling arguments of online analysis (Borodin and El-Yaniv, 1998); here it measures how tightly an entry-only controller can straddle a price signal it never observes. The fixed-price benchmark in part (ii) is the policy the too-high-toll objection actually proposes; a burn an operator resets by hand as demand shifts is itself an adaptive rule, and a coarser one than the automatic schedule. Two asymmetries make the guarantee conservative in the direction Sybil-resistance cares about. The burn errs upward: between registrations it sits inside $[c^*, m\,c^*)$, averaging $c^*/\ln m$ over a decay cycle, about $1.44\,c^*$ at $m = 2$, so its typical error is to over-price entry slightly rather than to leave the leader cloneable, and it is below the clone-resistant level only for the single block before a corrective registration. And a surge is absorbed in logarithmically many steps.

Corollary 1 (Bounded surge exposure)**.** *If demand jumps up by a factor R within one block, the burn regains the band $[c^*, m\,c^*)$ within $\lceil \log_m R \rceil$ registrations and admits at most $\lceil \log_m R \rceil$ clones before doing so. Exposure to a demand shock is logarithmic in its size, not linear.*

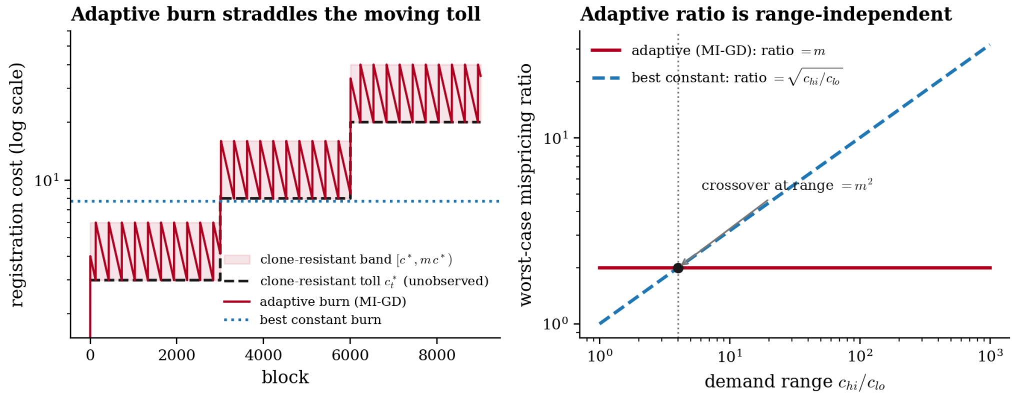

Figure 1: The adaptive burn tracks an unobserved, time-varying clone-resistant toll. Left: a multiplicative-increase, geometric-decay burn (increase factor $m = 2$, with an illustrative half-life chosen so the sawtooth is visible) stays inside the band $[c^*, m\,c^*)$, catching a rise in a few registrations and probing downward when entry pauses, while the worst-case-optimal constant under-prices at high demand and over-prices at low. Right: worst-case mispricing against demand range. The adaptive ratio is the increase factor m regardless of range (Theorem 2(i)); the best constant grows as $\sqrt{c_{\mathrm{hi}}/c_{\mathrm{lo}}}$ (part (ii)), so beyond a range of m^2 the adaptive rule strictly dominates. Both panels are computed from the exact dynamics.

The sawtooth of Figure 4, posted price climbing on bursts of entry and decaying when entry pauses, is this controller running on live demand; Figure 1 shows the same dynamics against a moving toll and the worst-case ratio the theorem bounds.

Table 1 reports the deployed parameters of the most-demanded subnet against the network defaults. Both adjustable knobs are set in the Sybil-resistant direction: the multiplicative step is larger and the decay is roughly forty times slower than the defaults, which is the correct response to the high copying pressure documented in Section 8. We do not claim that the precise multiplier and half-life are uniquely optimal. We claim two things

Table 1: Deployed registration-cost parameters versus network defaults.

Parameter	Most-demanded subnet	Network default
Multiplicative step per registration	2.0×	1.26×
Decay half-life (blocks)	14,400	360
Minimum burn (TAO)	0.99	n/a
Maximum burn (TAO)	100	n/a
Seats (capacity)	64	n/a

that the analysis does establish: under unknown, time-varying demand the demand-responsive burn stays within a fixed factor of the moving clone-resistant optimum that no constant burn can match (Theorem 2), and, by Proposition 7 below, lowering the toll only deepens duplication at the top. If anything, the measured externality argues for a higher floor and slower decay, subject to the countervailing concern of not deterring genuine innovators, which we treat as a calibration question rather than a corner solution.

6. Why the gradient is interior, not winner-take-all

Theorem 1 shows that abandoning grading is one way to keep free entry. Why not take it, and run a winner-take-all subnet immune to cloning by construction? Because grading is what elicits effort from a heterogeneous field. The softmax temperature T indexes a family that runs from winner-take-all to uniform.

Proposition 6 (Interior temperature)**.** *In the softmax family* $g(p) = e^{p/T}$ *the* *top-strategy share is continuous and strictly decreasing in T, with*

$$\lim_{T\to 0} \frac{e^{p^*/T}}{\sum_j e^{p_j/T}} = 1 \quad \textit{(winner-take-all)}, \qquad \lim_{T\to\infty} \frac{e^{p^*/T}}{\sum_j e^{p_j/T}} = \frac{1}{n} \quad \textit{(uniform)}.$$

At both extremes the marginal return to performance vanishes: at $T \to 0$ all reward sits on the single top seat, so every other seat faces a zero return to effort and capturing one seat captures everything; at $T \to \infty$ reward is independent of performance. Only at an interior T is the performance gradient positive at every seat.

Proof. As $T \to 0$ the largest exponent dominates and the share tends to 1; as $T \to \infty$ all exponents tend to 1 and the share tends to $1/n$. Strict monotonicity in T for a top strategy above the field follows because raising T compresses the differences in g. At $T \to 0$ the entire reward sits on one seat, so the marginal return to performance at every other seat is zero; at $T \to \infty$ the share is $1/n$ regardless of performance. Hence only interior T gives a positive gradient at every seat. ■

That an interior gradient *maximizes* elicited effort for a heterogeneous field, and is not merely necessary for a positive one, is the contest-theoretic conclusion of Moldovanu and Sela (2001); Szymanski and Valletti (2005); Lazear and Rosen (1981); we take it as given and use Proposition 6 only to locate the deployed temperature in the interior.

Figure 2 plots the top-strategy share against T for a representative full board. The deployed temperature $T = 100$ places the top share at 0.29, an interior value far from both the winner-take-all share of 1 and the uniform share of $1/n$. The deployed mechanism is therefore squarely in the regime

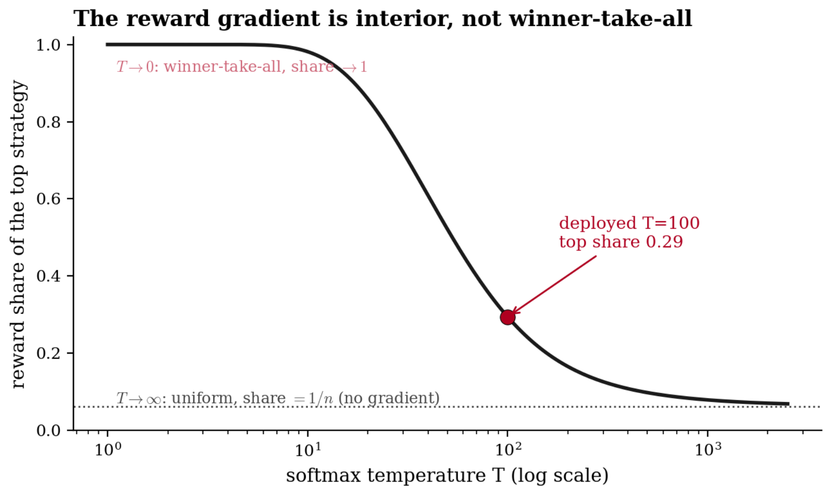

Figure 2: The reward gradient is interior. Top-strategy reward share as a function of the softmax temperature T for a representative sixteen-seat board (top rating normalized to zero, field spread to -500 Elo). Winner-take-all is the $T \to 0$ limit (share 1); uniform is the $T \to \infty$ limit (share $1/n$). The deployed $T = 100$ sits at an interior share of 0.29.

where Theorem 1 bites and the burn is indispensable: grading is retained for its effort benefits, which is exactly why free entry would be fatal.

7. The Toll-Reduction Paradox

We now study the comparative statics of a reduction in the burn, the change that the common objection requests. The result is that no reduction reduces cloning; it can only intensify the duplication of the highest-value strategies.

Proposition 7 (Toll-Reduction Paradox)**.** *Hold the mechanism fixed at a burn below the clone-resistant level $V\gamma$, the case any populated board occupies by Proposition 3, and reduce the burn further. Then the set of strategies a* *fresh copy can profitably clone weakly expands. A copy of a strategy with share w earns gross $V\,w/(1+w)$ and enters whenever $c < V\,w/(1+w)$, a threshold strictly increasing in w, so as the burn falls the strategies that become cloneable are those of progressively lower share, while the margin on copying the highest-share strategies, already cloneable, grows. Lowering the burn therefore deepens duplication exactly at the top of the board, where share, and so the value of copying, is greatest, and each copy it admits consumes a seat while adding no new strategy to the set D the network exists to produce. The registration cost is the only instrument that bounds this duplication, and reducing it can only loosen the bound.*

Proof. A fresh copy of a strategy with share w raises the normalizer to $Z(1+w)$ and earns $Vw/(1+w)$ (Proposition 1), so it enters iff $c < Vw/(1+w)$. The right-hand side is strictly increasing in w, so the cloneable set $\{w : Vw/(1+w) > c\}$ is an upper interval that expands monotonically as c falls, adding lower-share strategies while the surplus $Vw/(1+w) - c$ on every already-cloneable strategy rises; the largest such surplus is at the top share w^*. Each copy duplicates an existing strategy and so adds nothing to D (Definition 3) while occupying a seat. No registration-cost change other than a reduction enlarges the cloneable set, so the burn alone bounds duplication and lowering it weakly enlarges that set. ■

Corollary 2 (No accommodating adjustment)**.** *Fix the mechanism. The registration cost faces a trade-off with no interior escape: lowering it intensifies cloning (Proposition 7), while raising it rations honest participation (Proposition 3). In particular the too-high-cost objection cannot be met by lowering the cost without reviving the duplication of existing strategies that the cost* *exists to suppress. The adjustments that improve a dissatisfied participant's position lie outside the cost parameter: raise the strategy's performance, or enter a less contested venue.*

Proof. By Proposition 7 any reduction in the cost weakly enlarges the set of profitable clones, increasing duplication; by Proposition 3 any increase that improves clone-resistance prices out participation. The cost parameter thus moves the network along the frontier between resisting duplication and admitting entry and cannot step off it. Performance and venue choice, which lie outside the cost parameter, are the only levers that raise a participant's reward share without trading against either margin. ■

Corollary 2 gives the formal content of an otherwise informal dispute. The registration cost is not a free parameter that a complaint can move to everyone's benefit; it is the network's position on a frontier between resisting duplication and admitting entry, and the too-high-cost objection asks to move along it in the direction that revives the duplication the cost was raised to suppress.

8. Evidence

We bring four pieces of evidence from one deployed network. All quantities are computed from public on-chain data; no participant identities are used.

Registration cost is a demand thermometer. Figure 3 plots the registration cost of all 128 live subnets on a log scale, sorted. The mechanism is common to every subnet; the prices are not. In the June 2026 snapshot the most-demanded subnet prices entry at 10.9 TAO, rank one of 128, which is 5.8 times

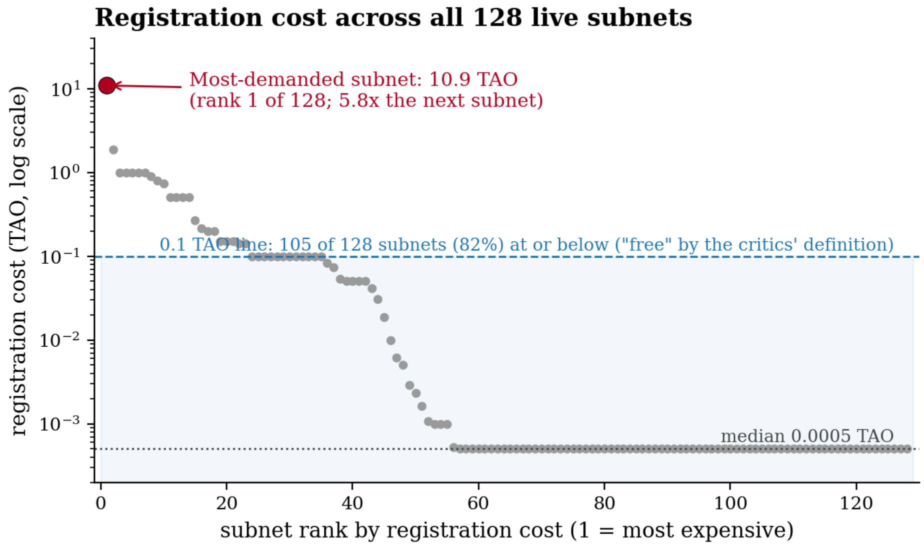

Figure 3: Registration cost across all 128 live subnets (log scale), sorted by cost. The most-demanded subnet is the rank-one outlier; 82 percent of subnets sit at or below the 0.1 TAO line commonly called free, because they attract little entry. Source: on-chain snapshot, June 2026.

the second-highest subnet (1.9 TAO) and roughly 22,000 times the median (0.0005 TAO); because the burn floats on a sawtooth, the multiple varies from day to day around this level, but the rank and the order of magnitude do not. Under the definition of free invoked by the objection, registration at or below 0.1 TAO, fully 105 of 128 subnets (82 percent) are already free, and 72 (56 percent) sit at the idle floor. They are cheap because almost no one is trying to enter them. The dispersion is therefore a dispersion in demand under a single mechanism, not a difference in rules. A high price on one subnet therefore reflects that it is the most in demand, not that its rule differs from the rest.

The price is market-set and self-correcting. Figure 4 shows the posted registration price over time on the most-demanded subnet. It climbed above 60 TAO in mid-May, decayed on its programmed half-life when entry paused, and has since settled into a sawtooth as entry continues. The posted price and the price actually transacted differ by construction: each registration doubles the posted figure, so what an entrant pays is about half of the price the *next* entrant then faces. In the recent window the amounts miners paid centered on a median near 12 TAO and ranged widely, from low single digits caught just after a decay to the low thirties at the May peak, while the posted price they bid against cycled through the high teens and into the high twenties. Nothing here is hand-set: the trajectory is a price discovering demand, rising on bursts and decaying toward the floor, exactly the adaptive toll whose worst-case optimality Theorem 2 establishes. Participants pay these sums voluntarily. Revealed preference weighs against the claim that the price is too high for those who enter: no one is required to enter, and entry demand persists at the highest price on the network. It is silent, by construction, on the honest entrants the burn rations out, the cost Proposition 3 identifies and that this cross-section does not measure.

The burn is what makes Sybil capture expensive. Figure 5 quantifies Propositions 2 and 4. Capturing all 64 seats by cloning the leading strategy costs about 6.4 TAO in the flat regime that a lowered burn implies (a sustained 0.1 TAO per seat, which requires neutering the adaptive response), versus roughly 6,200 TAO under the live doubling mechanism, where the price reaches the ceiling within a few registrations. The roughly thousand-fold gap is the Sybil-resistance the burn buys. Lowering the burn to that flat level would

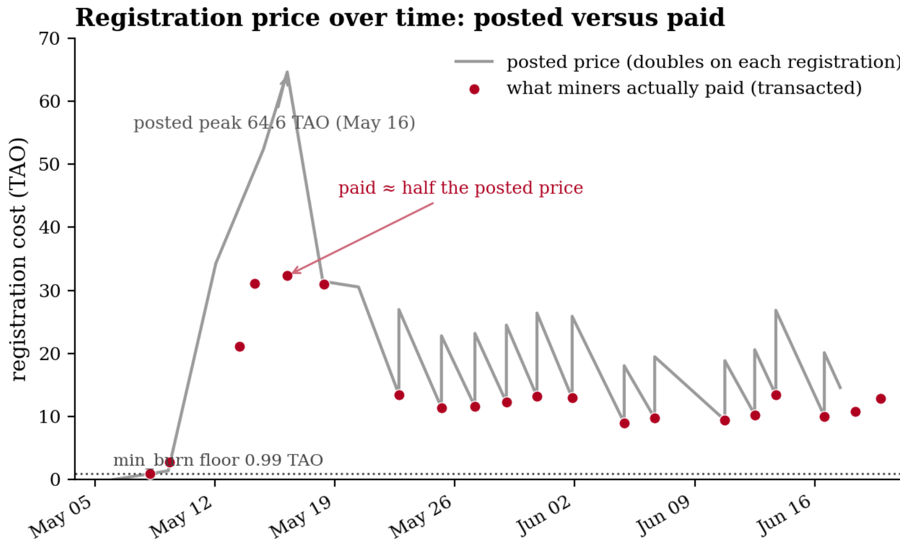

Figure 4: Subnet registration price over time on the most-demanded subnet. The grey line is the posted price (the recycle value, which doubles on each registration and decays on a 14,400-block half-life); the dots are what miners actually paid, anonymized. The mid-May posted peak exceeded 60 TAO; recent transactions cluster near 9 to 13 TAO, roughly half the posted price. The floor is 0.99 TAO. The price is set by demand against a fixed rule.

make capturing the entire subnet cost less than half the price of a single honest seat today.

Convergence in strategy space is observable even at the highest toll. Because every strategy is committed on chain as public plaintext, any competitor can read the current leader and copy it, so the cloning incentive of Proposition 1 acts on the entire field rather than on any single participant. Its signature is measurable. Treating each top strategy as a token set and taking pairwise Jaccard similarity, among the top nine commitments by rating on the most-demanded subnet six fall into a single connected near-identical cluster, where

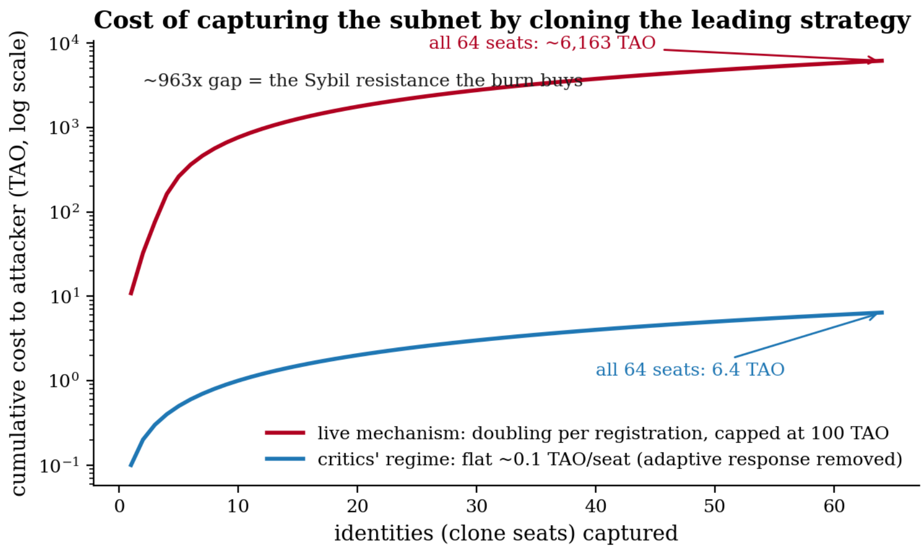

Figure 5: Cumulative cost to capture k seats by cloning the leading strategy (log scale). Under a flat 0.1 TAO price the full 64-seat capture costs 6.4 TAO; under the live doubling mechanism it costs roughly 6,200 TAO. The gap is the Sybil-resistance the burn provides.

an edge joins any pair sharing at least half its combined vocabulary. The cluster is robust to the threshold: six members for any cutoff from 0.4 to 0.6, five at 0.7, four at 0.8. The similarities separate cleanly, with every pair either at most 0.32 or at least 0.54 and nothing in between. Near-identical text is the signature of reproduction, not of convergent discovery: independent strategies that hit on the same idea express it in different words and would smear across the similarity range, whereas the clean bimodal gap, pairs either largely disjoint or sharing more than half their literal vocabulary, is what copying a public, readable leader produces. This holds although every one of these seats paid the highest entry price on the network, which is precisely Proposition 1 operating at the highest toll: the burn does not eliminate convergence, it bounds it. At zero price the bound vanishes and convergence becomes complete, the equilibrium Theorem 1 predicts under free entry with grading. That the field already clusters this tightly while paying the highest toll on the network is direct evidence that the dilution externality is large and that the toll, far from being too high, sits if anything below the level that would fully internalize it. The statistic is aggregate and identity-free, emitting no hotkeys or strategy text, and is reproducible from the deposited script.

9. Why a within-market experiment cannot settle it

A natural proposal is to test the claim by reserving a fraction of seats on the live subnet for free entry and comparing outcomes. This design is invalid. The object of interest, whether participants innovate or clone, is an equilibrium property of the whole market's entry cost, not a property of an individual seat. A cloner placed in a free seat copies the leading strategy that occupies a *paid* seat; the treatment leaks across units. In the language of causal inference this is a violation of the stable-unit-treatment-value assumption: one unit's treatment changes another unit's outcome, so no clean per-seat effect is identified. The within-market split also pays for the experiment in real diverted reward, since the free seats are filled by the very clones the design is meant to detect.

The correct unit of treatment is the market, not the seat. The clean design is a parallel subnet running identical code with free entry, a market-level treatment in which the natural worry that its participants would copy the paid subnet's strategies is not a confound but the dependent variable:

convergence to copies is the predicted outcome under free entry, so the measured object should be strategy diversity and the rate of genuinely new strategies, not absolute quality. The cheap subnets elsewhere on the network are not a substitute for this comparison: they run different tasks under different reward rules, several winner-take-all rather than graded, so their outcomes are silent on what free entry would do to a graded-reward debate market. The cross-subnet evidence of Section 8 is a price comparison, not an outcome comparison, and the part it relies on survives this heterogeneity: on every subnet, whatever its task or burn parameters, the price decays to its floor without entry, so a subnet resting at the floor reveals no entry pressure and one far above it reveals persistent demand. What the cross-section cannot recover, and what the same parallel-subnet design would, is the honest entry the current burn forecloses, the rationing margin of Proposition 3. A confounded cross-network comparison cannot settle the question; only a controlled one can. The companion paper (Anonymous, 2026) develops exactly that into a governance mechanism, in which a disputed parameter change is settled by a pre-registered experiment that both sides bond, so that the burden of proof is borne, symmetrically, by whoever asserts the counterfactual.

10. Conclusion

Permissionless reward networks face an impossibility, not a tuning problem. Free entry, Sybil-resistance, and graded reward cannot coexist; a network that wants the effort benefits of graded reward and the openness of permissionless entry must price entry, and the price that deters cloning is the marginal dilution externality a copy imposes on incumbents, a Pigouvian congestion toll best implemented in adaptive form: because the clone-resistant level moves with demand and is never observed, the demand-responsive rule that raises the burn on each entry and lets it decay otherwise stays within a fixed factor of the moving optimum in the worst case, a guarantee no constant burn attains. Charging that price in full has a cost the same analysis makes precise: because graded shares sum to one, the clone-resistant toll prices out all but a handful of seats, so resistance to cloning and a populated board cannot be maximized together. A live network sits at neither corner but at an interior toll that bounds cloning while keeping the board full. The objection that this toll is too high cannot, conditional on the mechanism, be met by lowering it: a populated board already sits below the clone-resistant level, so a cut cannot reduce cloning and instead deepens the duplication of the highest-value strategies, the direction the cross-network evidence shows is active even at the highest toll on the network. The same evidence shows a single common mechanism producing prices that span more than four orders of magnitude, with the highest price marking the strongest demand, not a harsher rule. The registration cost is thus a position on a frontier between resisting duplication and admitting entry, not a free parameter a complaint can move to everyone's benefit. What remains, once that is understood, is not a dispute over a number but a question about the principle of pricing entry at all, and that principle is what makes graded, permissionless reward feasible: in a network that pays for measured performance under open registration, a positive and adaptive entry toll is not a barrier to participation but the condition for rewarding quality at all.

References

Anonymous, 2026. Bonded dissent: A falsification protocol for parameter disputes in permissionless networks. Companion paper, authors withheld for review.

Babaioff, M., Dobzinski, S., Oren, S., Zohar, A., 2012. On bitcoin and red balloons, in: Proceedings of the 13th ACM Conference on Electronic Commerce (EC 2012), ACM. pp. 56–73. doi:10.1145/2229012.2229022.

Banerjee, A.V., 1992. A simple model of herd behavior. The Quarterly Journal of Economics 107, 797–817. doi:10.2307/2118364.

Baye, M.R., Kovenock, D., de Vries, C.G., 1996. The all-pay auction with complete information. Economic Theory 8, 291–305. doi:10.1007/BF01211819.

Bikhchandani, S., Hirshleifer, D., Welch, I., 1992. A theory of fads, fashion, custom, and cultural change as informational cascades. Journal of Political Economy 100, 992–1026. doi:10.1086/261849.

Borodin, A., El-Yaniv, R., 1998. Online Computation and Competitive Analysis. Cambridge University Press.

Chitra, T., Penna, P., Schneider, M., 2025. On Sybil-proofness in restaking networks. arXiv:2509.18338. Available at https://arxiv.org/abs/2509.18338.

Crawford, V.P., Sobel, J., 1982. Strategic information transmission. Econometrica 50, 1431–1451. doi:10.2307/1913390.

Douceur, J.R., 2002. The sybil attack, in: Druschel, P., Kaashoek, M.F., Rowstron, A.I.T. (Eds.), Peer-to-Peer Systems (IPTPS 2002), Springer, Berlin, Heidelberg. pp. 251–260. doi:10.1007/3-540-45748-8_24.

Emek, Y., Karidi, R., Tennenholtz, M., Zohar, A., 2011. Mechanisms for multi-level marketing, in: Proceedings of the 12th ACM Conference on Electronic Commerce (EC 2011), ACM. pp. 209–218. doi:10.1145/1993574.1993606.

Garimidi, P., Neuder, M., Roughgarden, T., 2025. Tullock contests in the wild: Applications in blockchains. SIGecom Exchanges .

Garimidi, P., Neuder, M., Roughgarden, T., 2026. Beyond winner-take-all procurement auctions. arXiv:2603.27779. Available at https://arxiv.org/abs/2603.27779.

Glazer, A., Hassin, R., 1988. Optimal contests. Economic Inquiry 26, 133–143. doi:10.1111/j.1465-7295.1988.tb01674.x.

Hirschman, A.O., 1970. Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States. Harvard University Press, Cambridge, MA.

Holmström, B., 1979. Moral hazard and observability. The Bell Journal of Economics 10, 74–91. doi:10.2307/3003320.

Janis, I.L., 1972. Victims of Groupthink: A Psychological Study of Foreign-Policy Decisions and Fiascoes. Houghton Mifflin, Boston.

Jensen, M.C., Meckling, W.H., 1976. Theory of the firm: Managerial behavior, agency costs and ownership structure. Journal of Financial Economics 3, 305–360. doi:10.1016/0304-405X(76)90026-X.

Karantias, K., Kiayias, A., Zindros, D., 2020. Proof-of-burn, in: Bonneau, J., Heninger, N. (Eds.), Financial Cryptography and Data Security (FC 2020), Springer. pp. 523–540. doi:10.1007/978-3-030-51280-4_28.

Lazear, E.P., Rosen, S., 1981. Rank-order tournaments as optimum labor contracts. Journal of Political Economy 89, 841–864. doi:10.1086/261010.

Lv, Y., Moscibroda, T., 2013. Fair and resilient incentive tree mechanisms, in: Proceedings of the 2013 ACM Symposium on Principles of Distributed Computing (PODC 2013), ACM. pp. 230–239. doi:10.1145/2484239.2484252.

Mazorra, B., Della Penna, N., 2023. The cost of Sybils, credible commitments, and false-name proof mechanisms. arXiv:2301.12813. Available at https://arxiv.org/abs/2301.12813.

Moldovanu, B., Sela, A., 2001. The optimal allocation of prizes in contests. American Economic Review 91, 542–558. doi:10.1257/aer.91.3.542.

Moulin, H., 2008. Proportional scheduling, split-proofness, and merge-proofness. Games and Economic Behavior 63, 567–587. doi:10.1016/j.geb.2006.09.002.

Olson, M., 1965. The Logic of Collective Action: Public Goods and the Theory of Groups. Harvard University Press, Cambridge, MA.

Pan, M., Mazorra, B., Schlegel, C., Mamageishvili, A., 2024. On Sybil-proof mechanisms. arXiv:2407.14485. Available at https://arxiv.org/abs/2407.14485.

Pigou, A.C., 1920. The Economics of Welfare. Macmillan, London.

Platt, M., Platt, D., McBurney, P., 2024. The Sybil attack vulnerability trilemma. International Journal of Parallel, Emergent and Distributed Systems 39, 446–460. doi:10.1080/17445760.2024.2352740.

Saleh, F., 2021. Blockchain without waste: Proof-of-stake. The Review of Financial Studies 34, 1156–1190. doi:10.1093/rfs/hhaa075.

Sisak, D., 2009. Multiple-prize contests: The optimal allocation of prizes. Journal of Economic Surveys 23, 82–114. doi:10.1111/j.1467-6419.2008.00557.x.

Spence, M., 1973. Job market signaling. The Quarterly Journal of Economics 87, 355–374. doi:10.2307/1882010.

Szymanski, S., Valletti, T.M., 2005. Incentive effects of second prizes. European Journal of Political Economy 21, 467–481. doi:10.1016/j.ejpoleco.2004.07.002.

Tullock, G., 1980. Efficient rent seeking, in: Buchanan, J.M., Tollison, R.D., Tullock, G. (Eds.), Toward a Theory of the Rent-Seeking Society. Texas A&M University Press, College Station, pp. 97–112.

Vickrey, W.S., 1969. Congestion theory and transport investment. American Economic Review 59, 251–260.

Wagman, L., Conitzer, V., 2008. Optimal false-name-proof voting rules with costly voting, in: Proceedings of the 23rd AAAI Conference on Artificial Intelligence (AAAI 2008), AAAI Press. pp. 190–195.

Appendix A. Numerical example for Figure 2

The representative board has sixteen seats with Elo ratings, top normalized to zero, at

$$\{0, -40, -75, -110, -140, -170, -200, -230, \\ -260, -290, -320, -350, -380, -410, -450, -500\}.$$

At $T = 100$ the top share is $e^0/\sum_j e^{p_j/100} = 0.294$; at $T \to 0$ it tends to 1; at $T \to \infty$ it tends to $1/16 = 0.0625$. The clone gains of Proposition 1 at $w^* = 0.294$ are $\Delta = w^*(1 - w^*)/(1 + w^*) \approx 0.160$ for an incumbent who doubles up, a 55 percent rise over its prior controlled reward, and the larger $\gamma = w^*/(1 + w^*) \approx 0.227$ for a fresh copy of the leader, which sets the clone-resistant burn $c^* = V\gamma$; the bound $\lfloor 1/\gamma \rfloor = 4$ of Proposition 3 then caps the fully clone-resistant board at four of the sixty-four seats.

Appendix B. Data sources and reproduction

The cross-subnet snapshot (Figure 3) is the per-subnet neuron-registration cost for all 128 live subnets (netuids 1 through 128; the root subnet, netuid 0, hosts no miner-registration market and is excluded), retrieved June 2026. The burn time series (Figure 4) combines the early price trajectory and the registration prices actually paid, with participant identities removed. The capture-cost curves (Figure 5) are computed from the deployed parameters in Table 1: a flat 0.1 TAO per seat versus a starting burn of 10.85 TAO doubling per registration and capped at 100 TAO, summed over 64 seats. The strategy-convergence statistic of Section 8 is computed from the on-chain strategy commitments as the token Jaccard similarity of their texts, with near-identical clusters taken as connected components of the graph that joins two strategies sharing at least the stated fraction of their combined vocabulary; the deposited script reports aggregate cluster sizes and the separation gap only, and emits no identities or strategy text. The convergence statistic examines the nine highest-rated of the 26 seats whose commitment text and rating were both legible in the snapshot, the top nine being where the cloning incentive concentrates, and the reported cluster is robust to the similarity threshold as Section 8 states. All figures and statistics are reproducible from the deposited scripts.