A Repeated-Game Framework for Incentives in Decentralized Infrastructure Protocols

Mustafa Qazi

Volt Capital

Abstract. We introduce a repeated dynamic incentive framework for characterizing when "compliance," or full-effort honest service provision, is incentive compatible in Decentralized Physical Infrastructure Networks (DePIN). We model quality control in this setting as a repeated moral-hazard problem between the protocol and each provider, where compliance is enforced by both slashing posted collateral and the discounted threat of demotion to probation tiers on a reputation ladder. Our main contribution is the "deterrence ratio" Γ, the worst-case ratio of a deviation's cost saving to its marginal probability of detection. We find that if any profitable deviation does not increase the fail probability relative to compliance, then binary public-outcome protocols cannot deter it. When all profitable deviations have positive detection gaps, compliance is sequentially incentive compatible if and only if immediate slashing plus discounted reputation loss exceeds Γ. We use this condition to formulate a protocol design problem mapping service primitives to stake requirements, reward schedules, probation rules, and audit frequency.

Keywords: Decentralized physical infrastructure networks · Repeated games · Moral hazard · Reputation · Slashing

1 Introduction

Infrastructure systems typically have three major inputs in producing service output: *capital*, or the hardware used to produce the infrastructure service, *operational effort*, or the ongoing maintenance and provision of capital, and *coordination*, the routing, load-balancing, and pricing layers that unify disjointed production nodes. In centralized infrastructure models, these inputs are under the complete control of one entity. In the emerging space of Decentralized Physical Infrastructure Protocols, or DePIN, protocols outsource the first two inputs by incentivizing individual *providers* to bring their own capital and provision service to the aggregating coordination layer in exchange for protocol-specified rewards. While the benefits of this model include more dynamic supply and the eliminated need for large upfront capital investment, the lack of control over production nodes gives rise to several new problems. One of these problems, and the focus of this work, is the issue of quality control: how can the protocol guarantee that work being done by providers is (a) actually being done and (b) at or above some threshold of quality? For example, a prevalent issue in DePINs is *self-dealing*,

where providers purchase or spoof the purchase of their own service to profit from token emissions. Even more detrimental to the network is the practice of spoofing signals that resemble actual work done without doing any work at all. In practice, disincentivizing or blocking this behavior is highly dependent on the type of service being provisioned: spoofing location data is structurally different from spoofing AI inference workloads. As a result, incentive-compatible mechanisms cannot generally be designed for *all* DePINs.

Our work takes a different approach in light of this reality: we quantify the difficulty of enforcing full-effort provider service provisioning as the *deterrence ratio*, a quantity that captures the binding deviation and network state as the worst-case scenario. Our work builds on the single-shot model of Milionis et al., a more general study of incentive-compatible signal recovery when signals are manipulable. We take their signal-recovery framework as a foundation for the audit layer and study the dynamic enforcement problem that remains once reports are aggregated on the protocol side.

Current DePIN designs use ad-hoc mixes of posted collateral slashing, auditors, and provider reputation tracking to address quality control. Our work formalizes this problem as a repeated moral-hazard game between providers and a committed public protocol, deriving necessary and sufficient conditions under which compliance is sequentially incentive compatible. The repeated game framework is essential because it reflects the two enforcement channels available to protocol designers. In a one-shot setting, the only deterrent is immediate punishment, which must be sized to cover the worst-case deviation. In the repeated game, our model can leverage the threat of reduced future rewards. The repeated-game setting also better reflects the decisions and actions of providers in real-world protocols: providers condition their behavior on past outcomes and have different temptations in different network conditions. Providers facing high costs of compliant behaviors in one state may choose to deviate even if compliance is optimal on average. Our state-conditional analysis addresses this concern.

1.1 Related Work

Our work is most closely related to that of Milionis et al. [12], who introduce the first formal treatment of eliciting signals from self-interested sources that may manipulate signals adversarially. They characterize a *source identifiability* condition, a necessary and sufficient condition for a strictly truthful elicitation mechanism, constructed via a family of mechanisms in the peer-prediction literature [13]. Our separation condition (Def. 4) rests on this foundation. Their work is a generalized framework in a single-shot model, from which we take their elicitation guarantees and build a repeated-game framework on top, specific to a DePIN environment. Our contribution begins after their reporting layer, investigating when the resulting monitoring technology is strong enough to support dynamic provider incentives through slashing and reputation.

Our recursive incentive constraints are closely related to repeated games with imperfect public monitoring in the work of Abreu et al. [1] and Fudenberg

et al. [4]. These works show how public signals and continuation values can support incentive-compatible behavior in repeated games. We use a similar dynamic incentive logic but study a more specialized object. Our model's protocol commits to a public automaton ex ante, applying mechanical rewards, slashing, and reputation transitions as a function of public outcomes. This is motivated by the reality of on-chain DePIN protocols that typically work with similar structures for quality control. We therefore do not characterize the full set of perfect public equilibrium payoffs and instead characterize when compliance is sequentially incentive compatible for providers under a fixed protocol. Classical repeated principal-agent works such as Radner [16] and Green and Porter [5] are related conceptually as they study how noisy public signals trigger future punishments.

Our modeling at the stage game level is a repeated moral-hazard problem in the sense of Holmstrom [7], where effort is unobservable and can only be inferred from noisy signals. Repetition matters because the protocol can discipline current deviations through both immediate punishment and the threat of future demotion in reputation. Our key result, the deterrence ratio Γ (3.3) isolates the *state-conditional* cost saving from deviation against the marginal detection ability of the audit. Our use of a reputation channel is most analogous to the no-shirking condition of Shapiro and Stiglitz [17], in which the threat of losing a rent-bearing position deters shirking under imperfect monitoring. We generate rent through a reputation ladder with stake slashing-backed demotion to lower tiers.

While our work is theoretical, it endogenizes several shared primitives and features of typical DePIN protocols in practice. Filecoin [15] secures decentralized storage through Proof-of-Replication and Proof-of-Spacetime, where they require providers to pledge collateral to be slashed in the event of failed proofs. Similarly, Helium [6] secures wireless coverage through Proof-of-Coverage, where hotspots verify one another's claimed locations. Although these whitepapers establish cryptographic verification procedures, they do not analyze incentives to derive conditions for sustained compliant behavior. In practice, both of these networks have suffered substantial location-spoofing and self-dealing attacks, which our aim works to target and quantify.

2 Model

2.1 Players and Information Structure

Time is discrete, $t = 1, 2, \ldots$, with common discount factor $\delta \in (0, 1)$. The strategic players are a finite set of *providers* indexed $i \in \mathcal{N} = \{1, \ldots, N\}$. The protocol is a committed public automaton that maps public outcomes to rewards, slashing, and reputation transitions. We assume that these parameters are fixed by the protocol designer and do not change strategically during the game. We define the collection of stake requirements, slashing rules, rewards, reputation transition rules, and lockup rules as a committed protocol

$$\Pi = (S, \lambda, r, T, \mathcal{G}, \Lambda_{\text{lock}})$$

This object is a non-strategic, committed automaton. These components are introduced below formally. We fix Π for Sections 2–4 and study the designer's choice of Π in Section 5.

Auditors are modeled as a part of the monitoring technology. They observe provider signals and submit reports to the protocol. We assume truthful reports as generated by an incentive-compatible recovery mechanism similar to Milionis et al. [12]. We expand on this strong assumption in section 2.1.

In every period, provider i has a privately observed state x_t^i drawn independently from a distribution μ over a finite set $\mathcal{X} = \{x_1, x_2, \ldots, x_M\}$. The state captures exogenous conditions outside of the provider's control like network health, hardware health, service demand, and other similar factors.

Let the provider action space be a finite set $\mathcal{A}$. We denote $a^* \in \mathcal{A}$ as *compliance*, or honest full-effort service provision. All other elements of $\mathcal{A}$ represent deviations of any kind: shirking, self-dealing, signal manipulation, etc. Conditional on x_t^i, provider i chooses an action $a_t^i \in \mathcal{A}$.

Let $\mathcal{S}$ denote a finite *signal space*. The action and state jointly determine a distribution over a private signal s_t^i observed by auditors assigned to provider i. Providers can strategically influence which distribution over signals is realized through a *manipulation correspondence* L:

$$L : \mathcal{X} \times \mathcal{A} \rightrightarrows \Delta(\mathcal{S}) \tag{1}$$

where $L(x, a) \subseteq \Delta(\mathcal{S})$ is the set of feasible signal distributions provider i can induce with action a and state x. This formalism is taken from Milionis et al., where the feasible set of application-specific signals (reports) is denoted similarly.

Assumption 1 (Manipulation Structure). *For each $x \in \mathcal{X}$, let $f_{a^*}(\cdot \mid x) \in \Delta(\mathcal{S})$ denote the signal distribution under compliance.*

(a) Under compliance, the signal distribution is a singleton: $L(x, a^*) = \{f_{a^*}(\cdot \mid x)\}$ [1].

(b) For each $(x, a) \in \mathcal{X} \times \mathcal{A}$, the set $L(x, a)$ is nonempty, compact, and convex in $\Delta(\mathcal{S})$.

(c) $f_{a^*}(s \mid x) > 0$ for all $x \in \mathcal{X}$ and for all $s \in \mathcal{S}$.

These assumptions are generally mild for real-world DePIN protocols. (a) ensures that compliant providers cannot manipulate their signal distribution. (b) ensures that optima exist when providers optimize over manipulation strategies and that randomization over different manipulation strategies is absorbed without loss of generality. (c) states that the compliance distribution has full support on $\mathcal{S}$ for each x, preventing trivial detection for signals impossible under compliance.

[1] Assumption 1(a) can be relaxed to allow compliant providers to optimize their signal presentation. The separation condition (Definition 4) then requires that deviating providers are caught more often than compliant providers even when both sides optimize their signal presentation. The remainder of the theory is mildly changed under this weaker formulation; the singleton case is a convenient special case that we adopt throughout.

Audit Layer and Public Outcomes Provider signal distributions are observed through the audit layer. Auditors receive signals generated by the provider's action and state, and an incentive-compatible reporting mechanism elicits reports from them. Following Milionis et al. [12], this mechanism can be thought of as a signal-recovery procedure for unverifiable information.

Conditional on the audit layer satisfying the truthfulness condition, the protocol aggregates reports and public randomness into a public outcome $z_t^i \in \mathcal{Z}$, where $\mathcal{Z}$ is a finite set with $|\mathcal{Z}| \geq 2$. The rest of the paper takes this induced public outcome mechanism as a given. In the baseline binary case $\mathcal{Z} = \{0, 1\}$, $z = 1$ denotes a 'fail' and $z = 0$ denotes 'pass'. Formally, the aggregation rule is a mapping

$$\mathcal{G} : \mathcal{S}^k \times \Omega \to \mathcal{Z} \tag{2}$$

where k is the number of auditor reports and Ω is a source of public randomness. With truthful reports $(s_1, \ldots, s_k)$ and randomness ω, the public outcome is $z = \mathcal{G}(s_1, \ldots, s_k, \omega)$. Our analysis applies to the binary outcome structure $\mathcal{Z} = \{0, 1\}$ throughout.

2.2 Public State and Player Histories

A convenient feature of repeated games on blockchains is the ability to define and publicly track player metadata. In classical repeated principal-agent games where the principal tracks a state [16], punishments are strategic and require sequential rationality. The punishing party must find it optimal to carry out the punishments.

In our setting, smart contracts track and enforce these punishments, eliminating the credibility problem entirely. Slashing and reputation updates execute as protocol code conditional on public outcomes.

Stake Requirements Upon entry into the protocol, each provider must post a *minimum stake collateral* $\underline{S} \in (0, \bar{S}]$, where $\bar{S} < \infty$ is the maximum stake[2]. The stake is locked for $\Lambda_{\text{lock}} \geq 1$ periods after any withdrawal request. Upon a fail result ($z_t^i = 1$), the protocol slashes a fraction $\lambda \in (0, 1]$ of the stake. In the event of slashing, the provider is barred from service provision until they replenish stake to $\underline{S}$. We write $S = \underline{S}$ throughout and treat stake as a constant on the equilibrium path, since the protocol requires immediate replenishment after a slashing event.

Each provider is assigned a *reputation* ρ_t^i from a finite set of reputation tiers $\mathcal{R}_n = \{G, P_1, \ldots, P_n\}$, where G denotes good standing and $P_1, \ldots, P_n$ are probation tiers[3]. The parameter n is a protocol design choice representing the number of passed public outcomes required for reinstatement to good standing. Given the public outcome, reputation evolves deterministically by a transition function $T : \mathcal{R}_n \times \mathcal{Z} \to \mathcal{R}_n$ defined below:

[2] Our analysis is agnostic to the choice of numeraire for stake collateral.

[3] Stake replenishment is also required for access to probation rewards. No rewards are given for undercollateralized providers in any case.

$$T(\rho, z) = \begin{cases} G & \text{if } \rho = G \text{ and } z = 0 \\ P_{j+1} & \text{if } \rho = P_j,\ j < n, \text{ and } z = 0 \\ G & \text{if } \rho = P_n \text{ and } z = 0 \\ P_1 & \text{if } z = 1 \end{cases} \tag{3}$$

The reward function (defined below) assigns $r(G) = r_G$ and $r(P_j) = r_j$ for $j = 1, \ldots, n$ with $r_1 \leq r_2 \leq \cdots \leq r_n \leq r_G$. Protocol rewards depend only on the reputation tier, not on the service state or action. Provider i's public history at time t is then:

$$h_t^i = (\rho_1^i, z_1^i, \ldots, \rho_{t-1}^i, z_{t-1}^i, \rho_t^i)$$

Timing and the Committed Protocol We now describe the timing of play under a fixed protocol Π.

In each period, the provider's tier ρ_t is public. The private state x_t is drawn and the provider chooses a_t and a manipulation from $L(x_t, a_t)$. The audit layer observes signals generated by the induced distribution and elicits reports from auditors. Conditional on the audit mechanism being truthful, the protocol aggregates reports into the outcome z_t. Rewards $r(\rho_t)$ are paid and slashing and probation transitions execute when fail outcomes are reported. Exit and re-entry decisions are made at the period boundary before x_{t+1}.

2.3 Payoffs

Given our players and information structures, we now model stage payoffs. Let the per-period service reward be $r : \mathcal{R}_n \to \mathbb{R}_+$ and the cost[4] of action a in state x be $c : \mathcal{A} \times \mathcal{X} \to \mathbb{R}_+$ where $c(a^*, x) > 0 \ \forall\, x$. The stage payoff for provider i in period t is:

$$u_t^i = r(\rho_t^i) - c(a_t^i, x_t^i) - \lambda S \cdot \mathbb{1}[z_t^i = 1]$$

where $\mathbb{1}[z_t^i = 1]$ represents the indicator function and λS is the slashing penalty. In the repeated game, provider i maximizes:

$$\mathbb{E}\left[\sum_{t=1}^{\infty} \delta^{t-1} u_t^i\right]$$

We further formalize this object as the continuation value function in section 4.

Definition 1 (Compliance). *Fix a committed protocol Π. The protocol implements compliance if, after every public history summarized by ρ and private state x, the provider weakly prefers compliance a^* to every valid deviation given the continuation values induced by Π.*

[4] Note that the cost function implies that compliance *must* incur a positive cost, a mild assumption in most DePIN protocols.

The provider's incentive to deviate depends on both the current state and the nature of the deviation. Following the moral hazard literature (Holmstrom 1979), the key quantity is the cost saving from deviating over complying. For some deviation $d \in \mathcal{A} \setminus \{a^*\}$, we define the *state-conditional, deviation specific* gain:

$$\Delta_d(x) = c(a^*, x) - c(d, x)$$

for each d and state x. The set of *state-relevant profitable deviations* is:

$$\mathcal{D}(x) = \{d \in \mathcal{A} \setminus \{a^*\} \mid \Delta_d(x) > 0\} \tag{4}$$

Crucially, $\mathcal{D}(x)$ depends on the state x, implying that a deviation that saves cost in one state may not save costs in another. This incentive constraint must hold for each (d, x) pair with $\Delta_d(x) > 0$. We then require the following assumption:

Assumption 2 (Non-Trivial Incentive Problem). *There exist $x \in supp(\mu)$ and $d \in \mathcal{A} \setminus \{a^*\}$ with $\Delta_d(x) > 0$.*

3 State-Conditional Separation

3.1 Public Monitoring and State-Conditional Detection

With our model in place, we can now see that the protocol is separated into two layers. The audit layer observes provider signals and delivers reports to the protocol, conditional on the auditing mechanism being incentive-compatible. The enforcement layer then processes these reports and compresses them into a public outcome, after which it can apply the reward, slashing, and reputation logic. From the provider's perspective, only the second layer affects dynamic incentives. After choosing an action and manipulation, the provider faces a distribution over public outcomes, which determine current penalties and future reputation.

Our baseline case is binary public outcomes. The coarseness of the pass and fail outcomes should be seen as a deliberate choice. On-chain logic will typically not discern between different manipulation types or deviation types, which means that it cannot condition punishments on the type of deviation. This is loosely related to the *full-rank condition* of Fudenberg et al. [4], in which it is required that outcome distributions induced by different player actions be linearly independent, requiring as many punishments as actions. We do not impose this assumption because it is highly unrealistic for real-world DePIN protocols, as even the most ambitious protocol designer will fail to enumerate the set of actions available to providers when setting corresponding punishments.

The relevant monitoring question is therefore not whether the audit layer identifies an action, but whether each profitable deviation makes punishment sufficiently more likely than it is under compliance. If a deviation saves cost without increasing the probability of a fail outcome, then the protocol's binary punishment is ineffective. If in the worst case the deviation lowers the fail probability compared to compliance, then the provider both saves cost and

becomes less likely to be punished than compliant providers. This motivates our concept of a *detection gap*. For each private state x, compliance induces a fail probability $\eta(x)$. A profitable deviation $d \in \mathcal{D}(x)$, paired with the provider's best feasible manipulation, creates a worst-case probability. We formalize this in the sections below.

3.2 Separation and Detection

For each state $x \in \mathcal{X}$, define the compliance-induced distribution over public outcomes:

Definition 2 (State-Conditional Compliance Distribution). *For each $x \in \mathcal{X}$:*

$$P_0^x(z) := \Pr\big(Z = z \mid a^*, x\big) \quad \textit{for each } z \in \mathcal{Z}$$

where signals $(s_1, \ldots, s_k)$ are drawn i.i.d. from $f_{a^*}(\cdot \mid x)$ and $Z = \mathcal{G}(s_1, \ldots, s_k, \omega)$.

Now, for each profitable deviation $d \in \mathcal{D}(x)$, the provider selects the manipulation $f_d \in L(x, d)$ that minimizes the probability of a fail outcome. This worst-case scenario gives the following condition its adversarial robustness.

Definition 3 (State-Conditional Deviation Distribution). *For each $x \in \mathcal{X}$ and $d \in \mathcal{D}(x)$:*

$$P_d^x(1) \;=\; \inf_{f_d \in L(x,d)} \Pr\big(Z = 1 \mid f_d\big)$$

$$P_d^x(0) \;=\; 1 - P_d^x(1)$$

where signals are drawn i.i.d. from f_d and $Z = \mathcal{G}(s_1, \ldots, s_k, \omega)$.

Note that the infimum is attainable as $L(x, d)$ is compact (Assumption 1b) and $\Pr(Z = 1 \mid \cdot)$ is continuous in the signal distribution. The separation condition is imposed after the audit layer. In the terminology of [12], their source identifiability condition concerns whether the mechanism can recover the information. Our $\varepsilon_d(x) > 0$ condition concerns whether this recovered information has enough statistical power to dynamically induce compliant behavior. Truthful signal recovery is a prerequisite for the audit layer.

With these definitions, we can now formalize the *state-conditional audit separation* condition:

Definition 4 (State-Conditional Audit Separation). *Define the false positive rate $\eta(x) := P_0^x(1)$ and the detection gap $\varepsilon_d(x) := P_d^x(1) - \eta(x)$ for each $x \in \mathit{supp}(\mu)$ and $d \in \mathcal{D}(x)$. The audit mechanism satisfies state-conditional audit separation if $\varepsilon_d(x) > 0$ for all such pairs.*

Separation requires that deviation increases fail probability even under optimal manipulation. This rules out deviations where the provider simultaneously saves costs and reduces their measured failure rate. When such deviations exist, compliance cannot be implemented under binary public outcomes under any reward or collateral scheme.

Proposition 1 (Impossibility Without Separation). *Consider committed protocols in which the rewards, slashing, and continuation values condition only on the binary public outcome $z \in \{0,1\}$, where $z = 1$ weakly reduces the provider's current or continuation payoff relative to $z = 0$. If there exists $x \in \mathrm{supp}(\mu)$ and $d \in \mathcal{D}(x)$ such that $\varepsilon_d(x) \leq 0$, then no such binary public-outcome protocol implements compliance against deviation d in state x.*

Note that this proposition is specific to mechanisms conditioning on binary outcomes. Richer structures of $\mathcal{Z}$ or other methods of out-of-protocol verification may deter deviations that are otherwise undetectable under binary aggregation.

3.3 Deterrence Ratio

The separation condition gives us a detection gap $\varepsilon_d(x)$ for each state–deviation pair. The relevant incentive question is whether this gap is large enough relative to the temptation gap $\Delta_d(x)$. We formalize this tradeoff in the following ratio.

Definition 5 (Deterrence Ratio). *For each $x \in supp(\mu)$ and $d \in \mathcal{D}(x)$, the deterrence ratio is defined as:*

$$\gamma_d(x) := \frac{\Delta_d(x)}{\varepsilon_d(x)}$$

Definition 6 (Global Deterrence Ratio).

$$\Gamma := \max_{x \in supp(\mu), d \in \mathcal{D}(x)} \gamma_d(x) = \max_{x \in supp(\mu), d \in \mathcal{D}(x)} \frac{\Delta_d(x)}{\varepsilon_d(x)}$$

The advantage of the global deterrence ratio Γ is that it captures the binding deviation-state pair, or the combination of the two that is hardest to deter.

4 Incentive Constraints and Dynamic Implementation

Section 3 established that binary outcomes carry information in the form of detection gaps for state–deviation pairs and compliance, summarized in the global deterrence ratio Γ. We now analyze whether this information, combined with protocol enforcement tools of slashing and reputation, is sufficient to make compliance incentive compatible in general. We show that the answer reduces to a single inequality.

4.1 Main Incentive Inequality

Fix a period t and suppose state-conditional audit separation holds for the audit mechanism. Let $V : \mathcal{R}_n \to \mathbb{R}$ denote the continuation value under compliant play where $V(\rho)$ is the ex ante expected discounted payoff from the current period onward for a compliant provider in tier ρ. Define the *continuation punishment*

$\Phi_\rho = V(T(\rho, 0)) - V(T(\rho, 1))$ as the value lost from a fail outcome relative to a pass.

Lemma 1. *Fix a committed protocol Π and continuation values V. Compliance is sequentially incentive compatible against all deviations $d \in \mathcal{D}(x)$ if and only if, after every reputation tier ρ and private state x:*

$$\lambda S + \delta \min_{\rho \in \mathcal{R}_n} \Phi_\rho \geq \Gamma \tag{5}$$

Proof. Consider a provider in state x contemplating deviation $d \in \mathcal{D}(x)$. Under compliance, the provider pays cost $c(a^*, x)$ and faces false-positive detection probability $\eta(x)$. Under deviation d, the provider pays $c(d, x)$ and faces fail probability greater than or equal to $\eta(x) + \varepsilon_d(x)$. The expected payoff loss from deviating rather than complying is:

$$\begin{aligned} &[c(d, x) - c(a^*, x)] + [\eta(x) + \varepsilon_d(x) - \eta(x)] \cdot [\lambda S + \delta \Phi_\rho] \\ &\qquad = -\Delta_d(x) + \varepsilon_d(x) \cdot [\lambda S + \delta \Phi_\rho] \end{aligned} \tag{6}$$

Deviation is then deterred if and only if (6) is nonnegative. Rearranging and dividing by $\varepsilon_d(x)$:

$$\lambda S + \delta \Phi_\rho \geq \frac{\Delta_d(x)}{\varepsilon_d(x)} \quad \forall x \in \operatorname{supp}(\mu),\ \forall d \in \mathcal{D}(x)$$

Since the left side does not depend on d or x, the pointwise condition holds for binding pairs if and only if it holds for the maximizing pair. Since we require this to hold at every tier, the binding constraint is the tier with the smallest Φ_ρ. Taking the minimum over tiers and maximum over state–deviation pairs:

$$\lambda S + \delta \min_{\rho \in \mathcal{R}_n} \Phi_\rho \geq \Gamma$$

4.2 The One-Shot Game

In the one-shot game, providers have no reputation or future values. A provider in state x chooses compliance a^* or a deviation d by comparing:

$$[r - c(a^*, x) - \eta(x)\lambda S] - [r - c(d, x) - (\eta(x) + \varepsilon_d(x))\lambda S] = -\Delta_d(x) + \varepsilon_d(x)\lambda S$$

Setting this difference nonnegative and rearranging gives $\lambda S \geq \Delta_d(x)/\varepsilon_d(x) = \gamma_d(x)$. Taking the max over all binding pairs, compliance is preferred if and only if $\lambda S \geq \Gamma$, which is exactly (5) with $\delta\Phi = 0$, or the staking-dominated regime.

This leaves protocol designers with just collateral to enforce compliance, which can grow massively when Γ is large and deviations are subtle and hard to detect. This motivates our study of the repeated game, where reputation can significantly reduce or completely substitute the economic cost of required collateral, leaving protocol designers with more tools to incentivize compliance when detection is difficult.

4.3 Continuation Punishment

Our main incentive inequality (5) depends on $\min_\rho \Phi_\rho$, which is determined by the reputation structure and reward schedule. We now derive this quantity from the primitives of our model.

Let $\bar{c} = \sum_x \mu(x) \cdot c(a^*, x)$ denote the expected cost of compliance and $\eta = \sum_x \mu(x)\eta(x)$ denote the average false positive rate. Under compliant play, the continuation values satisfy the following Bellman system:

$$\begin{aligned} V_G &= r_G - \bar{c} - \eta\lambda S + \delta[(1-\eta)V_G + \eta V_{P_1}] \\ V_{P_j} &= r_j - \bar{c} - \eta\lambda S + \delta[(1-\eta)V_{P_{j+1}} + \eta V_{P_1}] \quad \text{for } j < n \\ V_{P_n} &= r_n - \bar{c} - \eta\lambda S + \delta[(1-\eta)V_G + \eta V_{P_1}] \end{aligned} \tag{7}$$

These equations are computed under compliant play, where the provider chooses a^* in every state. Since the provider has not observed future x when the continuation value is evaluated, and x_t is i.i.d., the expected per-period cost and false positive rates are $\bar{c}$ and η respectively. The term κS is omitted from the Bellman system. This constant is borne in every period regardless of tier or action, shifting $V(\rho)$ by the same amount and canceling out.

Lemma 2 (Existence and Uniqueness of Continuation Values). *The Bellman system* (7) *has a unique solution* $\{V(\rho)\}_{\rho \in \mathcal{R}_n}$ *with all finite values.*

Proof. The Bellman system (7) defines an operator $\mathcal{T} : \mathbb{R}^{|\mathcal{R}_n|} \to \mathbb{R}^{|\mathcal{R}_n|}$. For any two candidate value vectors V, V':

$$|(\mathcal{T}V)_\rho - (\mathcal{T}V')_\rho| = \delta|(1-\eta)(V_{T(\rho,0)} - V'_{T(\rho,0)}) + \eta(V_{P_1} - V'_{P_1})| \leq \delta\|V - V'\|_\infty$$

Thus $\mathcal{T}$ is a contraction with modulus $\delta < 1$ under the sup norm. The result follows from Banach's fixed point theorem.

Proposition 2 (Continuation Punishment with n Probation Tiers). *In the n-period probation system, suppose the tier-dependent rewards $r_G, r_1, \ldots, r_n$ are monotone. The binding continuation punishment is:*

$$\min_{\rho \in \mathcal{R}_n} \Phi_\rho = \Phi_{P_1} = \sum_{j=1}^{n-1} \alpha^{j-1}(r_{j+1} - r_j) + \alpha^{n-1}(r_G - r_n) \tag{8}$$

where $\alpha = \delta(1-\eta)$.

Proof. Since $T(\rho, 1) = P_1$ for all ρ, we have $\Phi_\rho = V(T(\rho, 0)) - V_{P_1}$. Monotone rewards and the Bellman system (7) imply $V_{P_1} \leq V_{P_2} \leq \cdots \leq V_{P_n} \leq V_G$, so Φ_ρ is minimized when $T(\rho, 0)$ is smallest. For $n \geq 2$, $T(P_1, 0) = P_2$ and the minimum is $\Phi_{P_1} = V_{P_2} - V_{P_1}$; for $n = 1$, $T(P_1, 0) = G$ and $\Phi_{P_1} = V_G - V_{P_1}$. Evaluating the closed form below covers both cases.

To compute Φ_{P_1}, define $D_j = V_{P_{j+1}} - V_{P_j}$ for $j < n$ and $D_n = V_G - V_{P_n}$. Subtracting the two Bellman equations yields $D_n = r_G - r_n$. For $j < n$: $D_j = (r_{j+1} - r_j) + \alpha D_{j+1}$. Unrolling the recursion from $j = 1$ gives the result.

Corollary 1 (Two-Tier Reputation Case). *When $n = 1$, the reputation set becomes $\mathcal{R}_1 = \{G, P_1\}$ and $\Phi_{P_1} = \Phi_G = r_G - r_1$. Our incentive inequality (5) reduces to $\lambda S + \delta(r_G - r_1) \geq \Gamma$ and the minimum stake becomes $S = (\Gamma - \delta(r_G - r_1))/\lambda$.*

4.4 Participation Constraints and Deterring Sybil Re-entry

A provider who exits the protocol earns a per-period *outside option* $\bar{u} \geq 0$, representing the return from deploying their capital and hardware in alternative use. This value may be small in protocols with specialized equipment but will remain positive with our assumption that locked capital can always be deployed elsewhere.

Under compliance, a provider incurs expected false positive slashing cost of $\eta\lambda S$ per period and opportunity cost κS for locked capital, where $\kappa > 0$ is the per-period cost of locked capital. The participation constraint, or the condition under which a provider chooses to remain in the protocol, must also hold at the binding probation tier to give the threat of demotion credible force:

$$r_1 - \bar{c} - \eta\lambda S - \kappa S \geq \bar{u}$$

Rearranging, the minimum probation reward is then:

$$r_1 \geq \bar{u} + \bar{c} + S(\eta\lambda + \kappa) \tag{9}$$

Equation (9) is a sufficient flow condition. V omits the per-period capital cost κS, so we define $W(\rho) = V(\rho) - \kappa S/(1-\delta)$. The dynamic participation condition at tier ρ is $W(\rho) \geq \bar{u}/(1-\delta)$.

Assumption 3 (Cost of Re-entry). *Let $U_{lock} = \bar{u} \cdot \delta(1 - \delta^{\Lambda_{lock}})/(1-\delta)$ be the present value of forgone outside options during the lockup period. The cost of re-entry must satisfy:*

$$\kappa_{entry} + \lambda S + U_{lock} \geq \delta(V_G - V_{P_1})$$

where $\kappa_{entry} \geq 0$ is the cost of establishing a new identity[5]*.*

The left side is the total cost of re-entry: onboarding expenses, new additional capital requirements, and foregone outside option earnings during the lockup period while both stakes are committed. The right side represents the discounted benefit of escaping probation. Note that λ, S, and Λ_{lock} are protocol parameters and can always be set large enough to satisfy this assumption. The key tradeoff is that increasing these requirements will likely raise barriers for honest new providers. We formalize this optimization problem in Section 5.

[5] This assumption requires that each hardware node is limited to one on-chain identity, as creating two identities tied to the same hardware would allow trivial avoidance of probation. This assumption may be strong in certain DePIN verticals, but is typically handled with on-device onboarding keys or hardware attestation.

4.5 Compliance Incentive Compatibility

The terms in our main inequality (5) represent two different regimes of enforcement for the protocol designer. The first term, λS, is the immediate slashing penalty set by minimum stake requirements and slashing rates. The second term, $\delta \min_\rho \Phi_\rho$, is the discounted future cost of being caught with the threat of reputation punishments. Our main result shows that when these enforcement channels are both sufficient and providers find participation worthwhile, compliance is sequentially incentive compatible.

Theorem 1 (Compliance Under Slashing and Reputation). *Fix a committed protocol $\Pi = (S, \lambda, r, T, \mathcal{G}, \Lambda_{lock})$. Suppose Assumptions 1-3 hold, the audit layer satisfies state-conditional separation, and participation and re-entry are satisfied. Then compliance is sequentially incentive compatible against every profitable deviation $d \in \mathcal{D}(x)$ after every public reputation tier and private state if and only if*

$$\lambda S + \delta \min_{\rho \in \mathcal{R}_n} \Phi_\rho \geq \Gamma$$

Proof. By Lemma 2, the Bellman system (7) has a unique bounded solution V, the continuation values induced by the committed protocol under compliant play. Since the reputation transitions are deterministic, any public history is summarized by the reputation tier ρ.

Fix any tier ρ, private state x, and profitable deviation $d \in \mathcal{D}(x)$. Under compliance, the fail probability is $\eta(x)$. Under deviation and optimal manipulation, the fail probability is $\eta(x) + \varepsilon_d(x)$. The deviation saves cost $\Delta_d(x)$, but the expected punishment increases by

$$\varepsilon_d(x)\big[\lambda S + \delta \Phi_\rho\big]$$

Deviation is unprofitable if and only if

$$\lambda S + \delta \Phi_\rho \geq \frac{\Delta_d(x)}{\varepsilon_d(x)}$$

We require this condition for every ρ, x, and $d \in \mathcal{D}(x)$, which is equivalent to

$$\lambda S + \delta \min_{\rho \in \mathcal{R}_n} \Phi_\rho \geq \Gamma$$

Remark 1. The Bellman system (7) is linear in $|\mathcal{R}_n|$ unknowns with a unique solution. For any practical number of tiers, the system is solvable by standard methods.

5 Protocol Design

Theorem 1 establishes that a committed protocol implements compliance when the main inequality and all participation constraints are jointly satisfied. These

conditions involve protocol parameters that the designer controls: stake S, slashing rate λ, probation length n, and reward schedule $(r_1, \ldots, r_n, r_G)$. The parameters that fall (mostly) out of the protocol designer's control are the primitives determined by the service and state environments: deterrence ratio Γ, false positive rate η, expected compliance cost $\bar{c}$, and outside option $\bar{u}$. The designer's problem is to choose the controllable parameters to satisfy all constraints given the uncontrollable parameters, at minimum cost. We now formalize this optimization and introduce an applied framework for reward and audit design.

5.1 Audit Frequency and Redundancy

Before stating the designer's problem formally, we characterize how audit frequency and redundancy affect the separation condition and hence the deterrence ratio Γ. Suppose the protocol audits each provider independently with probability $p \in (0, 1]$ per period. When no audit occurs, the public outcome defaults to pass ($z = 0$). The effective false positive rate and detection gap scale accordingly. This scaling assumes the provider does not observe whether the current period will be audited before choosing a_t.

Lemma 3 (Audit Frequency). *Under random auditing with probability p, the effective false positive rate and detection gap are:*

$$\eta_{\mathit{eff}}(x) = p \cdot \eta_1(x), \qquad \varepsilon_{\mathit{eff},d}(x) = p \cdot \varepsilon_{1,d}(x)$$

where $\eta_1(x)$ and $\varepsilon_{1,d}(x)$ are the rates under certain auditing $p = 1$. The deterrence ratio under audit frequency p is:

$$\Gamma(p) = \max_{x,d} \frac{\Delta_d(x)}{p \cdot \varepsilon_{1,d}(x)} = \frac{\Gamma_1}{p}$$

where Γ_1 is the deterrence ratio under certain auditing.

Proof. With probability $1 - p$, no audit occurs and $z = 0$ regardless of the provider's action. With probability p, the audit proceeds with fail probability $\eta_1(x)$ under compliance or at least $\eta_1(x)+\varepsilon_{1,d}(x)$ under deviation. The unconditional fail probability is therefore $p \cdot \eta_1(x)$ under compliance and at least $p \cdot (\eta_1(x) + \varepsilon_{1,d}(x))$ under deviation. The detection gap is $p \cdot \varepsilon_{1,d}(x)$, and since $\Delta_d(x)$ is unaffected by audit frequency, the deterrence ratio scales as Γ_1/p.

Reducing audit frequency inflates the deterrence ratio and hence the required collateral. The minimum stake under audit frequency p is $S^*(p) = \Gamma_1/(p\lambda)$ in the staking-dominated regime. The designer trades off direct audit cost against collateral requirements.

5.2 Reward Design

The structure of (8) gives the protocol designer insight into how to optimally set rewards for their choice of n tiers. The binding continuation punishment is

expressed as a discounted sum of the reward increments along the ladder. Taking $\Delta_j = r_{j+1} - r_j$ for $j < n$ and $\Delta_n = r_G - r_n$, it is clear that the increments telescope towards the total reward span $\sum_{j=1}^{n} \Delta_j = r_G - r_1$. The geometrically declining weights in this equation then imply front-loading as the optimal reward scheme.

Proposition 3 (Front-Loaded Rewards). *Fix the floor r_1 and ceiling r_G. For any monotone schedule $r_1 \leq \cdots \leq r_n \leq r_G$, the continuation punishment Φ_{P_1} is maximized by the front-loaded schedule $r_2 = r_3 = \cdots = r_n = r_G$ where $\Phi_{P_1} = r_G - r_1$.*

Could a protocol designer instead set uniform rewards for all probation tiers? In this scenario, a provider at P_1 sees almost no benefit from passing a single audit, since they merely advance to P_2 at the same reward. The incentive to comply during probation then weakens with each additional probation tier. This increases the minimum reward spread and may push towards infeasible reward designs.

Remark 2 (Uniform Probation Rewards). Under uniform probation rewards $r_j = r_P$ for all $j \leq n$, all intermediate increments vanish and $\Phi_{P_1} = \alpha^{n-1}(r_G - r_P)$, which is decreasing in n. Any probation length $n \geq 2$ with uniform rewards will be counterproductive at the bounding tier. Setting a uniformly low probation reward and a high good-tier reward is strictly dominated by a front-loaded schedule in which the steepest gradient is between P_1 and P_2.

5.3 The Protocol Designer's Problem

We now formalize the optimization problem from the perspective of the protocol designer. The inputs to this problem are elements of the service environment, which must be measured or estimated for a given protocol's monitoring capability and workload type. These elements are the deterrence ratio under certain auditing Γ, the per-audit false-positive rate η_1, the expected compliance cost $\bar{c}$, the provider outside option $\bar{u}$, the per-period cost of locked capital κ, the discount factor δ, the per-audit cost c_{audit}, the maximum feasible stake $\bar{S}$. The outputs are controllable protocol parameters: probation period length n, audit frequency p, stake S, and the reward schedule $(r_1, r_2, \ldots, r_n, r_G)$.

We fix $\lambda = 1$ throughout this section. Our incentive constraint depends on slashing only through the product λS, while the cost of capital κS depends on S alone. Thus the designer always prefers maximal slashing with the smallest stake meeting the constraint.

Before stating the problem formally, we must derive the steady-state cost of operating the reputation system. Under compliance, the reputation tier evolves as a Markov chain on $\mathcal{R}_n$: a fail with probability η sends the provider to state P_1, and a pass with probability $1 - \eta$ sends the provider one tier toward G. The chain is both irreducible and aperiodic, so it admits a unique stationary distribution.

Proposition 4 (Steady-State Distribution). *Under compliance, the stationary distribution over tiers is:*

$$\pi_G = (1-\eta)^n, \qquad \pi_{P_j} = \eta(1-\eta)^{j-1} \quad \textit{for } j = 1, \dots, n$$

Proof. The balance equations give $\pi_{P_1} = \eta(\pi_G + \sum_{j=1}^{n} \pi_{P_j}) = \eta$ and $\pi_{P_j} = (1-\eta)\pi_{P_{j-1}}$ for $j \geq 2$, so $\pi_{P_j} = \eta(1-\eta)^{j-1}$. Then $\pi_G = 1 - \eta \sum_{j=1}^{n}(1-\eta)^{j-1} = 1 - (1-(1-\eta)^n) = (1-\eta)^n$.

A provider spends $(1-\eta)^n$ of time in good standing, which decreases exponentially in probation depth. Note that deeper probation does not increase the maximum achievable by Φ_{P_1} by Proposition 3. Instead, it expands the designer's ability to lower reward cost with more intermediate tiers below r_G, at the cost of keeping providers longer at those tiers. The expected per-period reward paid to a compliant provider in steady-state is then:

$$\bar{R}(n, \mathbf{r}) = (1-\eta)^n r_G + \eta \sum_{j=1}^{n} (1-\eta)^{j-1} r_j$$

Under audit frequency p, the effective false positive rate is then $\eta = p\eta_1$ and the deterrence ratio is $\Gamma(p) = \Gamma_1/p$ by Lemma 3. The designer then solves:

$$\begin{aligned} \min_{n,\, p,\, S,\, \mathbf{r}} \quad & \bar{R}(n, \mathbf{r}) + p \cdot c_{\text{audit}} \\ \text{subject to} \quad & S + \delta\, \Phi_{P_1}(n, \mathbf{r}) \geq \Gamma(p) && \text{(IC)} \\ & r_1 \geq \bar{u} + \bar{c} + S(\eta + \kappa) && \text{(Participation)} \\ & r_1 \leq r_2 \leq \cdots \leq r_n \leq r_G && \text{(Monotonicity)} \\ & p \in (0, 1], \quad S \in (0, \bar{S}], \quad n \in \mathbb{Z}_+ \end{aligned} \tag{10}$$

where $\Phi_{P_1}(n, \mathbf{r})$ is given by (8) with $\alpha = \delta(1-\eta)$. We omit the re-entry condition (3) as its left side is increasing in Λ_{lock}, which contributes no cost to the provider objective. The designer then satisfies it by setting the lockup period sufficiently long after solving the other parameters.

The program's structure is tractable and directly computable. For a fixed probation depth n and audit frequency p, the effective rates η and $\Gamma(p)$ are fixed and the remaining problem $(S, \mathbf{r})$ is a linear program. The objective $\bar{R}$ is linear in rewards, Φ_{P_1} is linear in the increments between rewards, and the participation and monotonicity constraints are linear in $(S, \mathbf{r})$.

Proposition 5 (Optimum Binding Constraints). *Assuming the monotonicity constraints are slack at the optimum, at any solution of* (10)*, the IC and participation constraints bind.*

Proof. By way of contradiction, suppose that IC is slack at some solution. Then r_G can be reduced by some small increment without violating constraints, strictly reducing $\bar{R}$, which contradicts optimality. Suppose similarly that participation is slack at some solution. Then, r_1 can be reduced, strictly reducing $\bar{R}$, a contradiction.

At the optimum, the designer then equates the marginal cost of collateral enforcement with the marginal cost of the reward spread. The designer chooses audit frequency to balance direct audit costs against the collateral and rewards required to cover $\Gamma(p)$.

Since this program is a linear program for each fixed (n, p), the global optimum is found by enumerating $n \in \{1, \ldots, n_{\max}\}$ over a grid of $p \in (0, 1]$ and solving each result. Furthermore, because the marginal deterrence benefit of additional reward tiers decays with α^{n-1}, setting $n_{\max}$ to be roughly 5 or less will suffice in practice. Solving the linear sub-programs can be done with standard methods and negligible computation.

5.4 Worked Example: Wireless Coverage DePIN

We consider a wireless-coverage protocol similar to Helium but with different hypothetical primitives for illustrative purposes. We denominate these primitives in cost units per epoch.

Symbol	Primitive	Value
Γ_1	Deterrence ratio under full auditing	20
η_1	Per-audit false positive rate	0.05
$\bar{c}$	Expected compliance cost	10
$\bar{u}$	Provider outside option	2
κ	Per-period cost of locked capital	0.10
δ	Discount factor	0.95
c_{audit}	Per-audit cost	1
$\bar{S}$	Maximum feasible stake	10

Table 1. Illustrative primitives for a wireless-coverage protocol

The most difficult of these primitives to estimate in practice is the deterrence ratio Γ, as it requires the protocol designer to consider the most profitable forms of non-compliance. For our wireless network, the ratio could correspond to running at reduced transmit power while asserting full effort, spoofing location signals, or responding selectively to workloads. For each deviation, estimating the difficulty of detection can be done with controlled experiments: red-team nodes can execute deviations and the rate of detection can be measured. We highlight that exact values are both unlikely to be derived in practice and unnecessary to apply our framework, as identifying the actual binding pair deviation is the most crucial element.

One possible outcome of this is that the binding deviation is simply undetectable under the current auditing scheme. If a deviating node can save cost while passing audits as often as an honest node, then $\varepsilon_{1,d}(x) \leq 0$ and separation fails. In this case, by Proposition 1, no setting of the remaining primitives can sustain compliance over time. This serves as a diagnostic test for the auditing mechanism that *must* be solved before further primitive calibration can continue.

In our example, as we will see, the combination of stake and future punishment can cover the required deterrence of $\Gamma_1 = 20$. Take the two-tier system ($n = 1$) where $\Phi_{P_1} = r_G - r_1$ and the stationary distribution is $\pi_G = 1 - \eta$, $\pi_{P_1} = \eta$. The expected reward cost is $\bar{R} = (1 - \eta)r_G + \eta r_1$. Substituting the binding participation and IC constraints, the marginal effect of stake on reward cost is:

$$\frac{d\bar{R}}{dS} = (\eta + \kappa) - \frac{1 - \eta}{\delta} = (0.15) - \frac{0.95}{0.95} = -0.85 < 0,$$

Thus, the optimum holds the maximum stake $S = \bar{S} = 10$.

$$r_1 = \bar{u} + \bar{c} + \bar{S}(\eta + \kappa) = 2 + 10 + 10(0.15) = 13.5,$$

$$r_G - r_1 = \frac{\Gamma_1 - \lambda\bar{S}}{\delta} = \frac{20 - 10}{0.95} \approx 10.53, \qquad r_G \approx 24.03.$$

As the stake is maxed at 10, the reputation spread of 10.53 covers the residual $\Gamma_1 - \lambda\bar{S} = 10$ that collateral does not cover. This quantifies the role of audit quality. If we suppose that an improved auditing scheme leads to a deterrence ratio of $\Gamma_1 = 10$, then $\lambda\bar{S} = \Gamma_1 = 10$. In this case, collateral alone suffices, freeing the protocol to only pay the participation floor instead of optimizing the reward spread.

References

1. Abreu, D., Pearce, D., Stacchetti, E.: Toward a theory of discounted repeated games with imperfect monitoring. Econometrica 58(5), 1041–1063 (1990)
2. Chiu, M.T.C., Mahajan, S., Ballandies, M.C., Kalabić, U.V.: DePIN: A framework for token-incentivized participatory sensing. In: 2025 IEEE International Conference on Blockchain and Cryptocurrency (ICBC), pp. 1–7. IEEE (2025). arXiv:2405.16495
3. Friedman, E.J., Resnick, P.: The social cost of cheap pseudonyms. Journal of Economics & Management Strategy 10(2), 173–199 (2001)
4. Fudenberg, D., Levine, D., Maskin, E.: The folk theorem with imperfect public information. Econometrica 62(5), 997–1039 (1994)
5. Green, E.J., Porter, R.H.: Noncooperative collusion under imperfect price information. Econometrica 52(1), 87–100 (1984)
6. Haleem, A., Allen, A., Thompson, A., Nijdam, M., Garg, R.: Helium: A decentralized wireless network. White paper (2018). `http://whitepaper.helium.com/`
7. Holmstrom, B.: Moral hazard and observability. Bell Journal of Economics 10(1), 74–91 (1979)
8. Holmstrom, B.: Moral hazard in teams. Bell Journal of Economics 13(2), 324–340 (1982)
9. Kandori, M.: Social norms and community enforcement. Review of Economic Studies 59(1), 63–80 (1992)
10. Lazear, E.P., Rosen, S.: Rank-order tournaments as optimum labor contracts. Journal of Political Economy 89(5), 841–864 (1981)
11. Lin, Z., Wang, T., Shi, L., Zhang, S., Cao, B.: Decentralized physical infrastructure network (DePIN): Challenges and opportunities. IEEE Network (2024)

12. Milionis, J., Ernstberger, J., Bonneau, J., Kominers, S.D., Roughgarden, T.: Incentive-compatible recovery from manipulated signals, with applications to decentralized physical infrastructure. arXiv:2503.07558 (2025)
13. Miller, N., Resnick, P., Zeckhauser, R.: Eliciting informative feedback: The peer-prediction method. Management Science 51(9), 1359–1373 (2005)
14. Pai, M.M., Roth, A., Ullman, J.: An anti-folk theorem for large repeated games with imperfect monitoring. arXiv:1402.2801 (2014)
15. Protocol Labs: Filecoin: A decentralized storage network. White paper (2017). `https://filecoin.io/filecoin.pdf`
16. Radner, R.: Repeated principal-agent games with discounting. Econometrica 53(5), 1173–1198 (1985)
17. Shapiro, C., Stiglitz, J.E.: Equilibrium unemployment as a worker discipline device. American Economic Review 74(3), 433–444 (1984)