Truthful by design: an incentive model for community verification of results-based conservation payments

Randy Sebastián Villalba Arango*

Whitepaper, version 1.3 — September 2026

Abstract

Tokenised carbon and biodiversity credits and on-chain impact funding promise to automate trust in nature-based solutions. Distributed ledgers, however, certify transactions, not what happens in a territory: they do not solve the oracle problem for nature, and current compensation schemes rarely give the communities that produce conservation outcomes a verifiable and fair role in measuring them. We propose a methodological framework for community-verified, results-based finance that can be coupled to a blockchain and tested in the field. Its core is a mechanism-design model in which the community paid for results also reports them, with three design levers: the weight of payment on self-reported outcomes versus verified activity records, the target of accountability audits, and community authority over territorial data. Absent future rents, verification capacity, not budget, caps what can be paid on self-reports; credible on-chain payment makes verification more necessary if and only if the sanctions a community faces outside the programme exceed the discounted cost of the conservation work it would stop doing if excluded. Implementable contracts are hybrid, audited networks of peer communities reconcile data authority with verifiability, and auditing representatives dominates auditing members when capture is a risk. On this basis we build a mixed monitoring and evaluation indicator framework, built on capture devices that communities own and govern, that keeps payment indicators apart from learning indicators, a negotiation and participation framework for funders, communities and verifiers, and a chain-agnostic architecture that places rules, payments and record proofs on-chain while raw territorial data stay under community custody. Colombia serves as a demanding reference case.

Keywords: nature-based solutions; community verification; results-based finance; mechanism design; oracle design; dMRV; DePIN; ReFi; monitoring and evaluation; data sovereignty; Web3

Key points

- Ledgers certify transactions, not territories: verification must be designed, not assumed
- Absent future rents, verification capacity, not budget, caps payment on self-reports
- On-chain payment needs more verification when outside sanctions dominate

*Verified Impact Protocol (VIP), Bogotá, Colombia. E-mail: randy.v@enreddados.org.

- Indicators that trigger payment must be kept apart from indicators used for learning
- Communities own and govern the capture devices; outside verification audits processes, not territories
- Rules, payments and proofs on-chain; raw territorial data under community custody

Summary for builders and funders

The problem. Nature-based finance increasingly pays for outcomes that its payees report themselves. Ledgers make payouts credible and records tamper-evident, but they do not solve the oracle problem for nature: nothing on-chain certifies that a forest is standing or that the reported work took place. Digital MRV adds satellites, sensors and registries (Belenky et al., 2022), and decentralised physical infrastructure networks (DePIN) crowdsource sensing through token incentives (The Block Research, 2026). Both inherit the same weakness whenever the party paid for the data also produces it.

The mechanism. We treat community verification as a mechanism-design problem with explicit trust assumptions. A community reports outcomes and activity records and is paid by a smart contract; verification, peer challenges and sanctions that play the role of slashing make truthful reporting incentive-compatible. The model supplies the bounds: how much payout can ride on self-reported data given the verification a funder can actually secure; when credible on-chain payment increases, rather than decreases, the verification needed; how audited, randomly assigned peer verification reconciles data sovereignty with verifiability; and how to prevent governance capture by community representatives.

The architecture. Chain-agnostic. Rules and escrowed payouts, hashed records, signed attestations and aggregated outcomes go on-chain; raw territorial data stay in community custody, on devices the community governs. The value is not the chain itself, since an escrow account also provides commitment, but the combination of commitment, tamper-evident records, auditable verification processes and data sovereignty. No token is required.

For funders. The framework turns the integrity risk of paying communities for self-reported results into explicit, testable design parameters. It is built for the channels that now direct money to communities, such as the Tropical Forest Forever Facility, which makes results-based payments to tropical forest countries and commits at least 20 percent of them to Indigenous peoples and local communities (World Bank Financial Intermediary Funds, 2026). Table 1 translates the model's terms into the vocabulary of cryptoeconomic design.

Table 1: Reading the model in Web3 terms

Model element	Web3 term	Meaning in the protocol
Truthful-reporting constraints (6), (7)	Incentive compatibility	Honest reporting pays more than misreporting, given detection and sanctions
Detection probabilities D_y, D_z	Verification or challenge probability	Chance that an inflated report or a fabricated record is caught
Sanction S (exclusion and external sanctions)	Slashing	What a reporter loses when caught
Future rents $K[\tilde{R}w(\beta) - c']$	Stake, value at risk	Value of staying in the programme, lost on exclusion
Commitment δ	Trustless escrow, credible commitment	Payouts execute by rule; the funder cannot renege
Record integrity ω	Tamper-evidence, timestamping	Records hashed at creation cannot be fabricated later
Peer flag probability ϕ	Decentralised oracle, peer prediction	Other communities challenge a report
Anonymous, non-reciprocal peer assignment	Random committee selection, anti-collusion	Peers cannot trade leniency
Capture κ, regard θ	Governance capture	A representative diverts payouts or acts against members
Data authority α	Data sovereignty, privacy by design	Raw territorial data stay with the community
Certified sensing ψ	Signed-at-source device (DePIN-style)	Readings that cannot be altered after capture
Learning indicators	Off-chain analytics	Measured to learn, never used to trigger payouts

1 Introduction

A growing set of Web3 instruments promises to channel capital to nature-based solutions and to the communities that sustain them: tokenised carbon and biodiversity credits, retroactive and quadratic funding rounds, impact certificates and self-executing payment contracts. The promise is that a public ledger can replace trust in intermediaries with verifiable records. It is only partly kept. Ledgers secure the traceability of transactions and help prevent double counting (Sorensen, 2023), but whether the data they record correspond to the world depends on the sources that feed them, an issue that has received little attention relative to the expectations placed on the technology (Caldarelli, 2020). The early record of tokenised carbon credits shows the gap: blockchain-based purchasers of Verra credits applied lower quality standards than conventional corporate buyers, and the registry later prohibited tokens based on retired credits (Payne Institute for Public Policy, 2022).

The compensation schemes underneath these tokens have integrity problems of their own. Evaluations of voluntary REDD+ projects in the Brazilian Amazon found that their crediting baselines overstated the deforestation they claimed to avoid (West et al., 2020). The communities whose territories produce the outcomes are often absent from verification, or present only as a formality. In 2024, Colombia's Constitutional Court found that the private entities behind a REDD+ project in the Pirá Paraná had not obtained genuine free, prior and informed consent from the Indigenous communities and had not effectively monitored the project's actual impacts. The validator and verifier had been contracted by the developers whose project it audited, and the Ministry of Environment told the Court that 49 percent of the 75 registered REDD+ projects it had reviewed faced complaints of alleged violations of social and environmental safeguards (Corte Constitucional de Colombia, 2024). Verification failed not for lack of technology, but because it was neither answerable to the communities nor robust to the incentives of those who were paid.

This paper proposes a methodological framework for community-verified, results-based finance of nature-based solutions and community work, designed to be coupled to a blockchain and tested in the field. Its premise is that communities should produce and govern much of the evidence on which they are paid. That premise creates an incentive problem that neither technology nor good intentions solve: when the community that is paid also reports the outcome that triggers payment, competence does not imply integrity. Community members can monitor forests as well as trained scientists (Danielsen et al., 2014), but monitoring by the monitored has proved weak where resources are at stake (Olken, 2007; Kahsay and Bulte, 2021). The same incentive applies to any party paid on results it reports itself, developers included. We focus on communities because shifting reporting to them is what fair governance requires, and because designing for truthful reporting protects them from the later invalidation of their credits.

At the core of the framework is an incentive model with three design levers: the weight of payment on self-reported outcomes relative to verified activity records, the target of accountability audits (the community's representative or its members), and the share of territorial data that remains under community authority. Ledgers enter the model through two parameters only, the credibility of payment commitment and the integrity of records, and not through the observability of territorial

outcomes. Without future rents, the payment that can safely depend on a self-reported outcome is capped by verification capacity whatever the budget. Our main result is a condition on primitives. Credible, self-executing payment raises both the reward for misreporting and the value of the future payments a community would lose if caught, so it makes verification more necessary if and only if the sanctions a community faces outside the programme exceed the discounted cost of the conservation work it would stop doing if excluded; otherwise, it substitutes for verification. Elite capture pushes toward complementarity. A cost-minimising funder pushes the weight on outcomes to the integrity bound, and a safety margin against uncertainty about detection has a computable price. Community data authority can be reconciled with verifiability through audited networks of peer communities, implementable contracts are generically hybrid, and auditing representatives dominates auditing members when capture is a risk and the community accepts the auditor.

On this basis we build three operational components. The first is a mixed monitoring and evaluation indicator framework that assigns every indicator a role in the protocol and keeps indicators that trigger payment apart from indicators used for learning. The second is a negotiation and participation framework that specifies which parameters funders, communities and verifiers set, who must consent to each and which bounds the model imposes, producing a term sheet for a pilot. The third is a chain-agnostic architecture that places rules, payments, record proofs and attestations on-chain and keeps raw territorial data under community custody. A pilot design, an evaluation plan and a sequence of implementation phases connect the framework to the field.

The framework is written for funders, networks and implementers who want their capital and infrastructure to support nature-based solutions with communities, and who need evidence that payments reward real outcomes without displacing the people who produce them. Large new channels are opening for such money: the Tropical Forest Forever Facility, launched at COP30, makes results-based payments to tropical forest countries on the basis of remotely sensed performance and commits at least 20 percent of them to Indigenous peoples and local communities (World Bank Financial Intermediary Funds, 2026). How that money is verified, shared and accounted for at community level is the problem this framework addresses. We use Colombia as a demanding reference case. It combines exceptional biodiversity and extensive collective territorial rights (in the Colombian Amazon alone, 231 legally recognised Indigenous reserves cover about 27 million hectares) with an advanced constitutional framework, and with armed conflict and unmet basic needs in regions where conservation projects operate (Corte Constitucional de Colombia, 2024). Requirements derived for this case are sufficient, though not always necessary, in simpler settings, and the model shows which of them relax where.

This is a methodological document. It specifies no payment levels, prices, remuneration or timelines: all values in the model and the figures are illustrative, and the values of an actual project are set after the diagnosis of Section 9, which combines an anthropological and social study of the territory with an analysis of its macroeconomic and social context.

Section 2 reviews what current mechanisms do and do not verify. Section 3 presents the model and Section 4 its results, which Section 5 illustrates numerically. Sections 6, 7 and 8 develop the

indicator framework, the negotiation framework and the blockchain architecture. Section 9 turns them into a pilot design, an evaluation plan and implementation phases. Section 10 states limitations and Section 11 concludes.

2 The verification gap

2.1 What current mechanisms verify

Certification standards place verification with third parties. Plan Vivo requires that at least 60 percent of the income from certificate sales return to communities, and relies on third-party validation and verification (Plan Vivo Foundation). The Climate, Community & Biodiversity Standards require participation and free, prior and informed consent, and more than 120 projects have been validated under them (Verra). Neither addresses verification when the community is both the recipient of payment and the source of the data that triggers it.

Web3 funding mechanisms solve different problems. Quadratic funding achieves first-best provision of public goods under its standard model, with variants designed to resist collusion (Buterin et al., 2019). Retroactive funding rewards impact after the fact, but the Optimism Collective acknowledged that the second round of its programme relied on qualitative descriptions without comparable impact indicators (Optimism Collective, 2023). Impact certificates such as Hypercerts allow several evaluators to assess the same claim, yet their promoters note that they are ineffective in low-trust environments without established evaluation norms (Gitcoin, 2026). Token-based governance adds its own risks: a study of 21 decentralised autonomous organisations found high concentration of voting rights and significant hidden costs (Feichtinger et al., 2024), and systemic bribery of voters is a realistic threat (Fábrega et al., 2025). These mechanisms allocate funds and aggregate votes; none verifies that a reported territorial outcome is true. Digital MRV systems combine remote sensing, connected sensors and ledgers and can connect to registries (Belenky et al., 2022); they automate measurement and reporting, but a system that pays the party that produces the data inherits the incentive problem this paper addresses.

2.2 What community monitoring can and cannot do

Competence is not the binding constraint. Across 34 tropical forest sites in four countries, community members and trained scientists produced similar estimates of the status and trends of species and resources (Danielsen et al., 2014). Integrity under payment is. In more than 600 Indonesian road projects, raising the probability of government audits from 4 to 100 percent reduced missing expenditures by eight percentage points, whereas grassroots monitoring had little average effect (Olken, 2007). Community monitoring works when the community watches a third party, as with primary health providers in Uganda (Björkman and Svensson, 2009), and is weaker when it must watch itself. In Ethiopian forest user groups, internal monitoring had no effect on livelihoods while top-down monitoring and punishment raised incomes (Kahsay and Bulte, 2021), and top-down monitoring of leaders drew in effort from members (Kahsay et al., 2023).

2.3 Non-extractive monitoring: communities that govern their own devices

Community monitoring can itself be extractive. Parachute science, in which outsiders take data and knowledge from communities they do not belong to and use the results to their own benefit, is extractive at its core, and avoiding it requires more than paying local people to collect data (Laumann et al., 2025). Research partnerships can siphon off data and experience while contributing little to local problems (Helm et al., 2023), and externally led community monitoring can relegate local actors to data collectors (Reyes-García et al., 2022). A payment protocol that relies on community monitors must therefore be designed so that monitoring serves the community, not only the funder.

Recent initiatives show that communities can govern the devices and data through which their environment is measured. Mapeo, an offline-first tool developed with Indigenous peoples, keeps data under local ownership through a peer-to-peer database without a central external server (Reyes-García et al., 2022). In London, a community programme distributed air-quality sensors to 60 community groups between 2021 and 2023 and let them choose where the sensors went, redistributing decision-making power; some group leaders moved from voluntary to paid roles, but imbalances persisted, as groups that asked for help received more of it (Schulte et al., 2025). In the Peruvian Amazon, a randomised trial of a programme that transferred satellite deforestation alerts to Indigenous communities and trained and paid community members to patrol in response found reduced forest loss in participating communities (Slough et al., 2021). Outside environmental monitoring, participatory guarantee systems offer a model of community-certified transparency: locally focused quality assurance systems that certify producers through the active participation of stakeholders, built on trust, social networks and knowledge exchange, and described as an alternative and a complement to third-party certification (May, 2019). The framework builds on these precedents: communities own and govern the capture devices, peers certify one another's monitoring processes, and outside verification examines documented processes rather than extracting territorial data.

2.4 What to pay for

Conditional payments for conservation can work, but the strongest evidence relies on external measurement. In a randomised trial in Uganda, tree cover declined by 4.2 percent in treatment villages against 9.1 percent in control villages, with outcomes measured by satellite (Jayachandran et al., 2017), while evidence across programmes remains of low or very low quality (Snilstveit et al., 2019). Whether to pay for outcomes or for inputs is open: paying for inputs has been proposed as a remedy for the integrity problem of community monitoring (Skutsch et al., 2014), and multitask theory explains why muted incentives can be optimal even when objective output measures exist (Holmström and Milgrom, 1991). Three principles frame the design problem: detection and punishment deter misconduct (Becker, 1968); future rents discipline those who could lose them (Shapiro and Stiglitz, 1984); and the more an indicator drives decisions, the more it is exposed to distortion (Campbell, 1979).

2.5 What fair governance requires

Community authority over data is a requirement, not an afterthought. The CARE principles ask that Indigenous Peoples' rights and interests in their data be recognised and their authority to control it strengthened (Carroll et al., 2020), and ledger-based frameworks for Indigenous data sovereignty already exist in genomic and health data (Mackey et al., 2022, 2026). The legitimacy of the verification chain matters as much as its accuracy, as the Colombian case of Section 1 shows (Corte Constitucional de Colombia, 2024). The gap is therefore specific. We are not aware of a framework that determines, for a community that both receives payment and reports the outcome that triggers it, which combination of payment object, audit target and data authority makes truthful reporting its best choice, how ledger-based commitment alters that combination, and how the answer translates into indicators, negotiation rules and an on-chain architecture.

3 Model

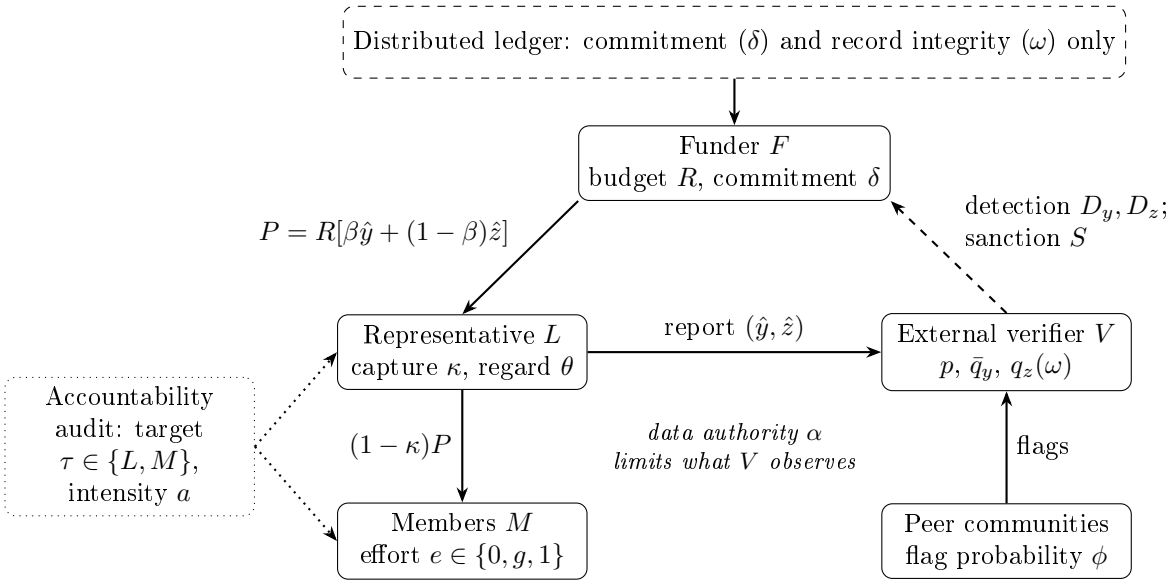


Figure 1: The protocol. Design levers: payment object β, accountability target τ and data authority α. Ledgers enter only through commitment and record integrity.

3.1 Environment

Figure 1 summarises the protocol. We study one funding period in which a funder (F) pays a community (C) for conservation outcomes. The community acts through a representative (L), who compiles and submits reports and receives and distributes payments, and through members (M), who perform the conservation work. We treat members as a single decision unit choosing collective effort $e \in \{0, g, 1\}$. Real conservation effort ($e = 1$) costs c. Paper compliance ($e = g$) consists of activities that are performed and documented but not targeted at the drivers of degradation, at cost

c_g with $0 < c_g < c$. Inaction ($e = 0$) costs nothing. The period outcome $y \in \{0, 1\}$, for example verified absence of forest loss, occurs with probability π_e, where $\pi_1 > \pi_g = \pi_0$. The difference $\Delta\pi = \pi_1 - \pi_0$ measures how informative the outcome is about effort; a low $\Delta\pi$ captures attribution problems such as exogenous drivers or leakage. Activity records $z \in \{0, 1\}$ exist if and only if $e \in \{g, 1\}$. Nothing in the analysis requires the reporter to be a community: L can be read as any party paid on outcomes it reports itself, including a project developer. We call it the community's representative because the protocol under study assigns reporting to communities.

Members also derive intrinsic and social motivation μ_τ from real effort, which lowers its net cost to $c - \mu_\tau$. We leave the sign of μ_τ open because the evidence points both ways, although crowding-out is the more frequent finding (Rode et al., 2015). In Ethiopia, external monitoring of leaders promoted forest conservation and drew in effort from members (Kahsay et al., 2023); in Colombia's Amazon Piedmont, collective payments strengthened social motivations to protect forests (Moros et al., 2019). The subscript τ indicates that motivation depends on the accountability regime defined below.

3.2 Design levers

The protocol has three levers. The first is the payment object. Payment depends on the reported outcome $\hat{y}$ and the reported activity record $\hat{z}$,

$$P = R\,[\beta\hat{y} + (1 - \beta)\hat{z}], \tag{1}$$

where $R > 0$ is the budget per period and $\beta \in [0, 1]$ is the weight on outcomes: $\beta = 1$ is pure payment by results, $\beta = 0$ pure payment for inputs, and intermediate values are hybrid contracts.

The second lever is authority over data. An external verifier inspects the aggregate report with probability p. Its ability to detect an inflated outcome report depends on what it can observe,

$$q_y(\alpha) = \bar{q}_y(1 - \alpha), \tag{2}$$

where $\alpha \in [0, 1]$ is the share of territorial signals held under community authority and not disclosed to outside verifiers, in line with the principle that Indigenous Peoples' authority to control their data should be recognised and strengthened (Carroll et al., 2020). In the baseline, data authority enters through this single channel; an extension lets α also raise data quality and legitimacy.

The third lever is the accountability target $\tau \in \{L, M\}$, audited with probability a. When $\tau = L$, the audit examines how the representative allocated the payment; it detects capture with probability q_c and imposes a personal sanction S_L. When $\tau = M$, the audit examines members' compliance with conservation rules; it detects non-conservation with probability q_m and imposes a sanction s_m. Because external oversight of customary authorities can itself infringe self-government, we let the motivational response to leader audits depend on the legitimacy ℓ of the auditor, that is, the degree to which the community accepts the auditor: $\mu_L = \mu_L(\ell)$, increasing in ℓ, so that an imposed auditor may produce $\mu_L < 0$.

The protocol may also include peer verification, in which neighbouring communities report on C's outcome under a scoring rule; $\phi \in [0,1]$ denotes the probability that an inflated report is flagged by peers. We take ϕ as a parameter in the main analysis. Its microfoundation (Supplementary Appendix B) rests on multi-task peer prediction, in which truthful reporting is the payoff-maximising equilibrium for binary signals (Dasgupta and Ghosh, 2013), avoiding the uninformative equilibria that affect single-task designs (Shnayder et al., 2016).

3.3 Records and commitment: the role of distributed ledgers

Distributed-ledger technology enters the model through exactly two parameters. The first is record integrity ω: records timestamped when they are created are harder to fabricate after the fact, so the probability of detecting fabricated activity records, $q_z(\omega)$, increases in ω. The second is commitment: without a self-executing rule, F honours payment only with probability $1-\delta$, whereas an escrowed smart contract sets $\delta = 0$. The effective budget is $\tilde{R} = (1-\delta)R$. Commitment is not specific to ledgers: an escrow account or a trust with legal enforcement also sets δ to zero, so every result about commitment below applies to any credible commitment device. What distinguishes ledgers in the model is that they combine commitment with record integrity. The ledger enters neither $\bar{q}_y$ nor α: it certifies when a record was written and that the payment rule executes, not that the territory is as reported. This modelling choice is consistent with the observation that the accounting advantages of ledger-based carbon assets come from transparency and resistance to double counting rather than from any capacity to observe the underlying asset (Sorensen, 2023), and with the oracle problem: whether the data entering a ledger match the world depends on the sources that supply them (Caldarelli, 2020). Ledgers can, however, be combined with sensors whose readings are signed at the source. We allow for this as an extension in which certified sensing raises outcome-detection precision to $\min\{1, \bar{q}_y + \psi\}$; Section 4 shows that it changes how much verification a contract requires but not the sign condition at the centre of our main result.

3.4 Timing and sanctions

The protocol $(\beta, p, \alpha, \tau, a, \phi, \omega, \delta)$ is public. Members choose e; nature draws y, and records exist if $e \geq g$. The representative submits $(\hat{y}, \hat{z})$ with $\hat{y} \geq y$ and $\hat{z} \geq z$, since only inflation is payoff-relevant, and peers file their reports. With probability p the report is verified and peer flags are resolved. If inflation is detected, the inflated component is withheld, the community is excluded from future rounds and it incurs external sanctions $S_{\text{ext}} \geq 0$ that do not depend on this funder's payments, such as legal liability, reputational loss or loss of certification with other buyers. The protocol is stationary and repeated, with discount factor $d \in [0,1)$ and outside option normalised to zero, so the sanction is

$$S(\tilde{R}, \beta) = S_{\text{ext}} + K\,[\tilde{R}\, w(\beta) - c'_\tau], \qquad K \equiv \frac{d}{1-d}, \quad w(\beta) \equiv \beta\pi_1 + (1-\beta), \quad c'_\tau \equiv c - \mu_\tau, \tag{3}$$

where the bracket is the community's per-period net surplus from participating, which we require to be non-negative. The fixed part of this sanction, $S_{\text{ext}} - Kc'_\tau$, can be negative, because exclusion also spares the community the cost of the conservation work. With a finite remaining horizon of T rounds, K is replaced by $K_T = d(1-d^T)/(1-d)$, the discounted number of future rounds lost on exclusion, and all results hold with K_T. The payment then executes with probability $1-\delta$, the representative retains a share $\kappa \in \{0, \bar{\kappa}\}$ and distributes the rest, and with probability a the accountability audit falls on τ. The representative's personal sanction S_L does not scale with $\tilde{R}$. This sequence adapts the costly-state-verification setting, in which an agent's realised outcome is observed only by the agent unless a verification cost is borne (Townsend, 1979).

3.5 The representative's objective

The representative decides what to report. It retains the share κ_τ of payments and places weight $\theta \in [0,1]$ on members' payoff, so that its effective stake in community money is $\lambda_\tau \equiv \kappa_\tau + \theta(1-\kappa_\tau)$. External sanctions are shared like payments, members bear the cost of conservation work, and the representative additionally bears a personal sanction $S_R \geq 0$ for detected misreporting, such as removal from office or legal liability, which does not scale with $\tilde{R}$. When $\lambda_\tau > 0$, dividing the representative's reporting constraints by λ_τ yields the constraints below with the effective sanction

$$S_\tau(\tilde{R}, \beta) = S_{\text{ext}} + \frac{S_R}{\lambda_\tau} + K\Big[\tilde{R}\, w(\beta) - \frac{\theta}{\lambda_\tau}\, c'_\tau\Big], \qquad A_\tau \equiv S_{\text{ext}} + \frac{S_R}{\lambda_\tau} - \frac{\theta}{\lambda_\tau}\, Kc'_\tau, \tag{4}$$

where A_τ is its fixed part; a representative with $\lambda_\tau = 0$ gains nothing from inflation. For a fully accountable representative ($\theta = 1$, $S_R = 0$), $\lambda_\tau = 1$ whatever κ_τ and $A_\tau = A^0_\tau \equiv S_{\text{ext}} - Kc'_\tau$, the benchmark we use to interpret the results. In reduced form, the elasticity of S_τ with respect to $\tilde{R}$ satisfies $\eta_S - 1 \gtrless 0 \iff A_\tau \lessgtr 0$.

3.6 Constraints

Throughout, S denotes the representative's effective sanction $S_\tau(\tilde{R}, \beta)$ in (4). The probabilities of detecting inflated outcome reports and fabricated records are

$$D_y = 1 - [1 - p\, q_y(\alpha)](1-\phi), \qquad D_z = p\, q_z(\omega). \tag{5}$$

Truthful outcome reporting requires that, when $y = 0$, the expected gain from inflation not exceed the expected sanction,

$$(1 - D_y)\, \beta \tilde{R} \leq D_y S, \tag{6}$$

and truthful record reporting requires, when $e = 0$,

$$(1 - D_z)(1-\beta)\tilde{R} \leq D_z S. \tag{7}$$

The representative refrains from capture if and only if

$$\bar{\kappa}\,\mathbb{E}[P] \leq a\,q_c S_L \cdot \mathbf{1}[\tau = L], \tag{8}$$

which determines the equilibrium share κ_τ. Anticipating κ_τ, members prefer real effort to paper compliance if and only if

$$(1-\kappa_\tau)\tilde{R}\beta\Delta\pi + a\,q_m s_m \mathbf{1}[\tau = M] \geq c - \mu_\tau - c_g, \tag{9}$$

and prefer real effort to inaction if and only if

$$(1-\kappa_\tau)\tilde{R}[\beta\Delta\pi + (1-\beta)] + a\,q_m s_m \mathbf{1}[\tau = M] \geq c - \mu_\tau. \tag{10}$$

A protocol implements truthful real conservation if and only if (6), (7), (9) and (10) hold together.

3.7 Maintained assumptions

All parties are risk neutral and liability is limited, so sanctions are bounded; the protocol is stationary, exclusion is permanent and there is no renegotiation. The representative's regard for members, θ, is exogenous, external sanctions are shared like payments, and the personal sanction S_R does not scale with $\tilde{R}$. Detected misreporting is sanctioned with exclusion, whereas shortfalls in effort are not: a community that reports truthfully after low effort simply receives less under a contract that pays for reported outcomes and activities, and member-level sanctions s_m are applied under the community's own rules rather than by the funder. Extending exclusion to detected non-compliance would add a term to (9) and (10) without changing the reporting constraints on which Proposition 1 rests. The outcome is one-dimensional; multidimensional social and environmental outcomes with dimensions of uneven measurability are the natural multitask extension, following the logic by which contracts may rely on fixed pay even when good objective output measures exist (Holmström and Milgrom, 1991). Peer verification is summarised by ϕ. Community monitoring is assumed to be technically competent, consistent with evidence that community members and trained scientists produced similar estimates across tropical forest sites (Danielsen et al., 2014); the model's question is integrity, not competence.

Table 2: Notation

Symbol	Meaning
β	Weight of payment on the self-reported outcome
p, $\bar{q}_y$, α	Report verification probability; outcome-detection precision at $\alpha = 0$; data authority
ϕ	Probability that peers flag an inflated outcome report
$q_z(\omega)$	Precision in detecting fabricated records, increasing in record integrity ω
δ, $\tilde{R}$	Commitment gap; effective budget $\tilde{R} = (1-\delta)R$
S, S_{ext}	Community sanction; external sanctions independent of this funder
d, K	Discount factor; $K = d/(1-d)$
c'_τ, A_τ, A^0_τ	Net cost of real effort $c - \mu_\tau$; fixed part of the effective sanction; its benchmark $S_{\text{ext}} - Kc'_\tau$
θ, λ_τ, S_R	Representative's weight on members; its effective stake $\kappa_\tau + \theta(1-\kappa_\tau)$; its personal sanction
τ, a	Accountability target (L or M) and audit probability
$\bar{\kappa}$, q_c, S_L	Maximal capture share; capture-detection precision; representative's sanction
q_m, s_m	Member-compliance detection precision and sanction
π_0, $\Delta\pi$	Outcome probability without real effort; informativeness of the outcome
c, c_g, μ_τ, ℓ	Cost of real effort; cost of paper compliance; motivational response; auditor legitimacy

4 Results

We characterise when a protocol implements truthful real conservation and how the three design levers and ledger-based commitment move that set. Throughout, $B \equiv \beta\tilde{R}$ denotes the outcome stake, the expected payment that rides on the self-reported outcome, and $X_\tau \equiv c - \mu_\tau - c_g - a\, q_m s_m \mathbf{1}[\tau = M]$ denotes the net cost advantage of paper compliance over real effort under regime τ. Proofs are in Supplementary Appendix A. Where a result evaluates S at a particular contract, we write S for $S(\tilde{R}, \beta)$ and state the value of β.

4.1 Self-verification

Lemma 1. *If $\beta > 0$ and $D_y = 0$, the community inflates its outcome report whenever $y = 0$.*

With no chance of detection, (6) reduces to $\beta\tilde{R} \leq 0$, which fails. Payment by results verified only by the payee is not incentive compatible, whatever the community's monitoring competence (Danielsen et al., 2014); what fails is integrity once payment depends on the report.

4.2 The outcome-stake band

Lemma 2. *Taking κ_τ as given, truthful outcome reporting and real rather than paper effort are jointly implementable if and only if*

$$\frac{X_\tau}{(1-\kappa_\tau)\Delta\pi} \leq B \leq \min\{B_{\max}, \tilde{R}\}, \qquad B_{\max} \equiv \frac{D_y\,(A_\tau + K\tilde{R})}{1 - D_y + D_y K(1-\pi_1)}. \tag{11}$$

(i) Without future rents ($K = 0$), $B_{\max} = D_y A_\tau/(1-D_y)$, which equals $D_y S_{\text{ext}}/(1-D_y)$ for a fully accountable representative, and the existence of the band does not depend on the budget. (ii) With future rents ($K > 0$), $B_{\max}$ increases in $\tilde{R}$.

The lower bound is the outcome stake needed to make real effort pay more than paper compliance; the upper bound is the largest stake verification can keep honest. Without future rents, the funder can rescale B through β but no amount of money raises the upper bound: money cannot close a verification deficit. With rents, money can, but only through the value the community attaches to staying in the programme, which requires long relationships and credible exclusion. The lemma formalises the warning that the more an indicator drives decisions, the more it is exposed to distortion (Campbell, 1979), and identifies future rents as the only budgetary counterweight.

4.3 The funder's choice of payment object

Lemma 3. *Suppose the funder takes the budget and verification as given and minimises expected payment $\tilde{R}\,w(\beta)$. If $\pi_1 < 1$ and some β is implementable, the optimal contract sets β at the largest implementable value.*

Paying on outcomes is cheaper than paying for activity records because outcomes may fail even under real effort. A cost-minimising funder therefore raises the weight on outcomes until an integrity or participation constraint binds. When the binding constraint is (6), the behaviour of $\beta_{\max}$ in Proposition 1(e) describes how the optimal contract responds to commitment. The design rules in Section 9 accordingly concern how far outcomes can be pushed, not whether funders would want to push them.

A contract set exactly at the integrity bound is fragile: if the funder overestimates detection, it induces misreporting. Because $\beta_{\max}$ is increasing in D_y whenever it is positive, robustness has a simple form.

Lemma 4. *Suppose the funder knows only that $D_y \geq \underline{D}$. Setting $\beta = \beta_{\max}(\underline{D})$ guarantees truthful outcome reporting for every $D_y \geq \underline{D}$. Relative to setting β at $\beta_{\max}(\hat{D})$ for a point estimate $\hat{D} > \underline{D}$, the guarantee costs $\tilde{R}(1-\pi_1)[\beta_{\max}(\hat{D}) - \beta_{\max}(\underline{D})]$ in expected payment. If $\beta_{\max}(\underline{D}) < X_\tau/[(1-\kappa_\tau)\tilde{R}\Delta\pi]$, no contract guarantees truthful real conservation over the whole range, and the funder must raise verification or accept the risk.*

The same logic applies to uncertainty about A_τ, since $\beta_{\max}$ is increasing in A_τ. The price of robustness is paid in expected payment, and it rises with the funder's uncertainty about the verification it actually commands.

4.4 Main result: when commitment and verification are complements

Proposition 1. *At a given contract β, an increase in payment commitment (a fall in δ, hence a rise in $\tilde{R}$): (a) raises the minimum outcome-detection probability required by (6) if $A_\tau > 0$ and lowers it if $A_\tau < 0$; (b) raises the minimum record-detection probability required by (7) if $A_\tau > 0$ and lowers it if $A_\tau < 0$; (c) strictly raises the accountability intensity $a\,q_c S_L$ required by (8) to deter capture when S_L does not depend on $\tilde{R}$; (d) relaxes (10). Moreover, (e) $\beta_{\max}$ falls if $A_\tau > 0$ and rises if $A_\tau < 0$, while $B_{\max}$ rises whenever $K > 0$.*

Credible payment raises both the reward for misreporting and the value of the future payments a community would lose if caught. Which effect dominates depends on the fixed part of the sanction, A_τ. For a fully accountable representative, commitment and verification are complements if the sanctions a community faces outside the programme exceed the discounted cost of the conservation work it would stop doing if excluded, and substitutes otherwise. Two readings follow. Where external legal, reputational or market sanctions are strong, a self-executing payment raises the stakes of misreporting faster than it raises the cost of exclusion, and must be accompanied by stronger verification. Where the programme is a community's main source of net surplus, because work is costly, intrinsic motivation is modest and the community values the future, credible payment disciplines through rents and partly substitutes for verification. A less intuitive implication is that high intrinsic motivation, which lowers c'_τ, pushes toward complementarity: when communities would conserve anyway, payments are largely rents whose credibility raises the temptation to inflate reports. This reading requires care. When $c'_\tau \leq 0$, members conserve without payment: $X_\tau < 0$, constraints (9) and (10) hold for every β, and payment buys no additional conservation. In that region the protocol does not create incentives but makes a transfer verifiable. Its justification then rests on communities' claims to benefit from conservation they already perform rather than on additionality, that is, on whether payment generates conservation beyond what would have happened anyway, which evaluations of voluntary REDD+ projects have found to be overstated (West et al., 2020); complementarity then concerns the integrity of transfers rather than of incentives. The most exposed configuration is therefore a credibly committed pure-results contract with weak outcome verification where $A_\tau > 0$, that is, a ledger that certifies the transaction but not the territory, in a setting where exclusion is not the main deterrent. Part (b) implies that, in the same region, commitment raises the returns to record integrity, so the two ledger functions should be deployed together.

Corollary 1. *If $c'_\tau \geq 0$, then $A_\tau - (S_{\text{ext}} - Kc'_\tau) = [S_R + Kc'_\tau\,\kappa_\tau(1-\theta)]/\lambda_\tau \geq 0$. Capture of the payment by a representative with $\theta < 1$, and personal sanctions on the representative, both push settings toward complementarity.*

A representative who keeps part of the payment and cares less about members than members do weighs the savings from exclusion, which accrue to members as avoided work, less than it weighs the payments at stake. Where elite capture is a concern, credible payment is therefore riskier. This links Proposition 1 to the accountability target in Proposition 4: deterring capture also reduces the verification that commitment requires.

Certified sensing. The sign conditions in Proposition 1 depend on A_τ and not on D_y. Raising outcome observability through certified sensing ($\psi > 0$) therefore relaxes (6) and raises $\beta_{\max}$ without changing whether commitment complements or substitutes for verification. Even if ledgers were bundled with sensing that relaxed our assumption that they do not improve observability, the main result would stand; what would change is how far outcome payments can be pushed. If sensor data are disclosed to outside verifiers, sensing also lowers effective data authority, which reintroduces the trade-off in Proposition 2.

4.5 Data authority and peer verification

Proposition 2. *Let S be evaluated at $\beta = 1$. (a) $\beta_{\max}$ is decreasing in α; without peers, a pure-results contract requires $p \geq \tilde{R}/[(\tilde{R}+S)\,\bar{q}_y(1-\alpha)]$, which becomes unattainable as $\alpha \to 1$. (b) If $\phi \geq \tilde{R}/(\tilde{R}+S)$, (6) holds for every β and every α. With k peers whose signals have accuracy r and a flag-triggered check of precision q_f, this requires $q_f > \tilde{R}/(\tilde{R}+S)$ and $k \geq \ln[1-\tilde{R}/((\tilde{R}+S)q_f)]/\ln(1-r)$, together with the incentive condition in Lemma B.2 and an audited network check (Lemma B.3).*

Community authority over territorial data and outside verifiability trade off, unless verification is relocated from outside auditors to networks of communities bound by reciprocal verification arrangements. Such relocation is compatible with data authority because signals stay within community networks, but it depends on the peer mechanism resisting uninformative agreement; for binary signals, a multi-task design makes truthful reporting the payoff-maximising equilibrium (Dasgupta and Ghosh, 2013). Supplementary Appendix B derives the flag probability and shows that anonymous, non-reciprocal assignment of peers removes the gain from reciprocal leniency. The check that a flag triggers is itself a report by communities, so the verification problem reappears one level up; it terminates only if that check is audited. Because the audit targets the documented check, a process record, rather than the territory, record integrity is what makes it feasible (Lemma B.3). Coalitions with side payments remain outside the analysis.

4.6 Hybrid contracts

Proposition 3. *If $X_\tau > 0$, no pure-input contract ($\beta = 0$) implements real effort. If $D_y < \tilde{R}/(\tilde{R}+S)$, with S evaluated at $\beta = 1$, no pure-results contract implements truthful reporting. When both hold and the band in (11) is non-empty, every implementing protocol is strictly hybrid, and the admissible interval of β narrows as α or κ_τ rise and as $\Delta\pi$ falls.*

This qualifies the proposal that paying for inputs rather than outputs would largely resolve the integrity problem of community monitoring (Skutsch et al., 2014): pure input payment works only when paper compliance is not the cheaper option. Generically, neither pure form implements truthful real conservation, echoing the multitask logic by which muted or mixed incentives can be optimal even when objective output measures exist (Holmström and Milgrom, 1991).

4.7 Accountability target

Proposition 4. *Real effort, constraint (9), is easier to sustain under leader-targeted than under member-targeted accountability if and only if*

$$(\bar{\kappa} - \kappa_L)\tilde{R}\beta\Delta\pi + \mu_L(\ell) - \mu_M \geq a\,q_m s_m. \tag{12}$$

Suppose capture is deterred under $\tau = L$, so $\kappa_L = 0$, and $\mu_L(\ell) \geq \mu_M$. Then leader-targeted accountability dominates unless direct deterrence of members is large. The motivational ranking is supported by evidence that top-down monitoring led leaders to work harder, which drew in effort

from members (Kahsay et al., 2023). When the auditor is imposed and ℓ is low, μ_L may fall below μ_M and the ranking can reverse; legitimacy is a design variable, not a courtesy. Propositions 1 and 4 also interact: commitment enlarges the first term of (12), raising the value of leader-targeted accountability, while raising the audit intensity needed to deter capture.

5 Numerical illustration

We normalise the nominal budget to $R = 1$, so that costs and sanctions are fractions of one period's payment. All values in this section are illustrative and correspond to no project; actual values are set per project after the diagnosis of Section 9. Table 3 reports the ranges of the planned numerical exploration. Where the literature offers an order of magnitude, we use it to bound a range, never to fix a value: audit probabilities in public programmes can be as low as 4 percent (Olken, 2007), and the effect of conservation effort on observed outcomes can be modest relative to background variation (Jayachandran et al., 2017). The full exploration will be run after the Delphi study of Section 9 informs the ranges of the contested parameters; it will report regions rather than shares, through two-parameter phase diagrams and a shallow classification tree fitted to a Latin hypercube sample, and its code includes unit tests that reproduce every closed-form result.

Table 3: Parameter ranges for the numerical exploration ($R = 1$)

Parameter	Range	Basis
β	grid on $[0, 1]$	Design lever
p	$[0, 0.5]$	Low audit rates (Olken, 2007)
$\bar{q}_y$	$[0.5, 1]$	Stylised
ψ (extension)	$[0, 0.3]$	Certified sensing
α	$[0, 1]$	Design lever
ϕ	$[0, 0.9]$	Stylised; Supplementary Appendix B
$q_z(\omega)$	$[0.3, 1]$	Stylised; rises with record integrity
δ	$[0, 0.5]$	Stylised
S_{ext}	$[0, 2]$	Spans both signs of A_τ; sign region informed by the Delphi study
d	$[0, 0.9]$	$K = d/(1-d) \in [0, 9]$
θ, S_R	$[0, 1]$, $[0, 2]$	Representative's regard for members; personal sanction
S_{ref}, η_S (isoelastic, robustness)	$[0.01, 2]$, $[0, 1.5]$	Reduced-form check
π_0, $\Delta\pi$	$[0, 0.5]$, $[0.02, 0.5]$	Modest effects (Jayachandran et al., 2017); $\pi_1 \leq 1$
c, c_g/c	$[0.1, 1]$, $[0.1, 0.9]$	Stylised
μ_L/c, μ_M/c	$[-0.3, 0.3]$	Sign left open
$\bar{\kappa}$	$[0, 0.5]$	Stylised
a; q_c, q_m	$[0, 0.5]$; $[0.3, 1]$	Stylised
S_L, s_m	$[0, 2]$, $[0, 0.5]$	Stylised

5.1 Illustrative results

We report illustrative results at two reference points with a fully accountable representative ($\theta = 1$, $S_R = 0$): $d = 0.5$ (so $K = 1$), $c = 0.3$, $c_g = 0.18$, $\mu_L = \mu_M = 0$, $\pi_0 = 0.3$, $\Delta\pi = 0.4$, $\delta = 0.2$,

$p = 0.3$, $\bar{q}_y = 0.8$, $q_z = 0.8$, $\psi = 0$, $\bar{\kappa} = 0.1$, $a = 0.3$, $q_c = 0.8$, $S_L = 1$, $q_m = 0.65$ and $s_m = 0.1$. The two points differ only in S_{ext} (0.8 and 0.1), which gives $A_\tau = 0.5$ and $A_\tau = -0.2$. They are chosen so that constraint (9) is feasible, which it is not at the midpoints of Table 3. The figures illustrate the propositions; they are not calibrated to any programme.

Figure 2 shows the (α, ϕ) plane. At both points, no contract is implementable without peer verification, and the peer flag probability required rises with data authority: from 0.04 to 0.27 when $A_\tau > 0$, and from 0.52 to 0.63 when $A_\tau < 0$, as α goes from zero to one. Where peer verification is strong, the funder's optimal contract pays for outcomes alone; near the boundary of the implementable region it is hybrid. The region is larger when $A_\tau > 0$ (81 percent of the plane) than when $A_\tau < 0$ (36 percent), because the second point has weaker external sanctions. Figure 3 displays Proposition 1: with $A_\tau > 0$, the outcome-detection probability required at a given contract rises with committed payment and $\beta_{\max}$ falls, while with $A_\tau < 0$ both move in the opposite direction. Figure 4 maps Proposition 4: auditing the representative eases the effort constraint when the motivational response to leader audits exceeds that to member audits, and over a wider range of motivational gaps when the risk of capture is higher.

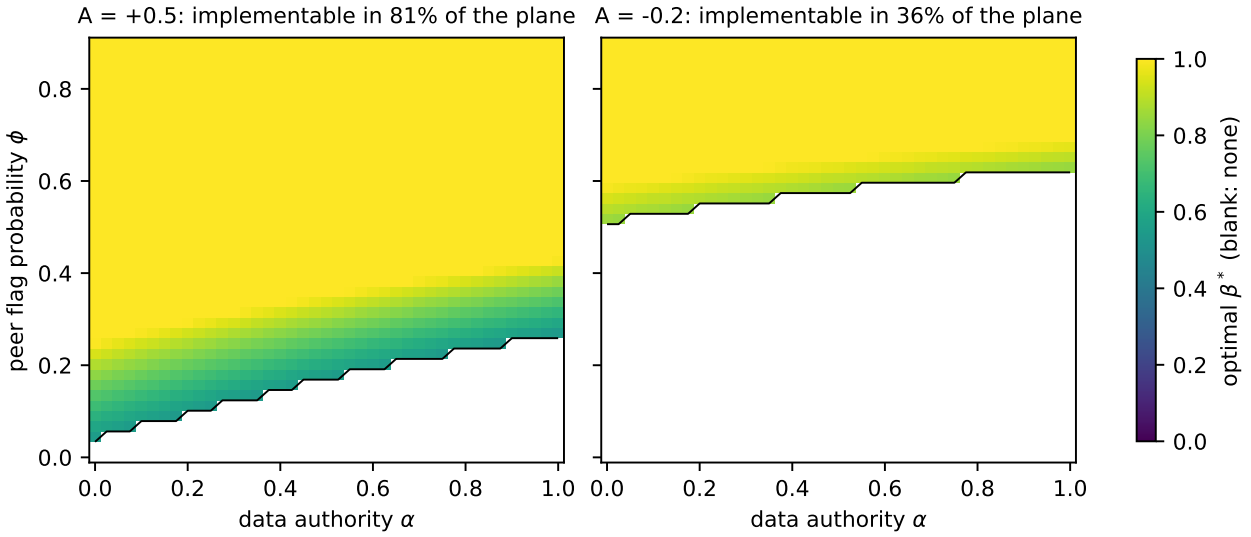


Figure 2: Implementability and the funder's optimal outcome weight β^* in the (α, ϕ) plane at the two reference points (leader-targeted accountability). Blank cells: no implementable contract. Illustrative, not calibrated.

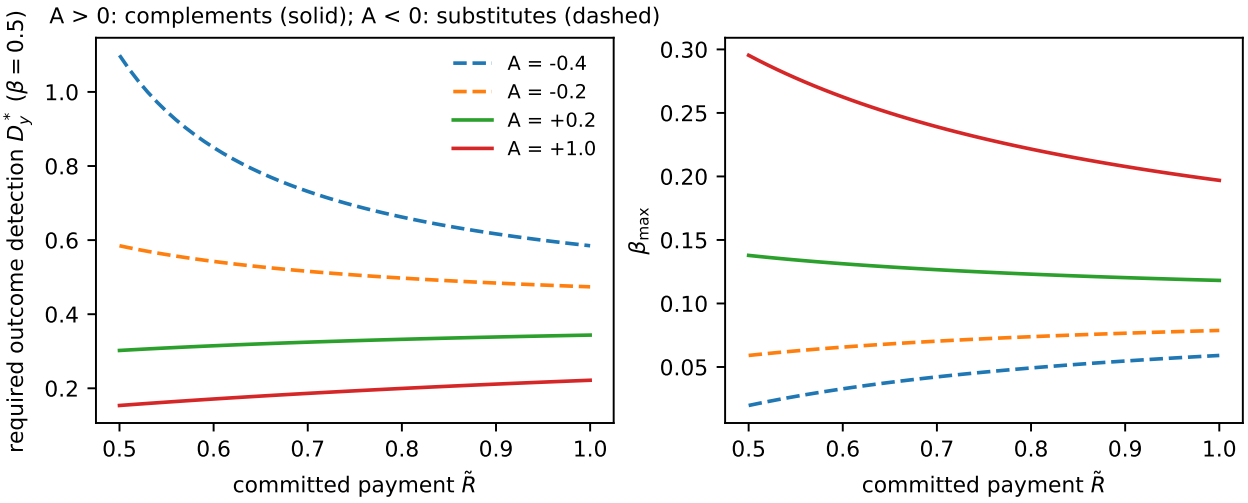


Figure 3: Proposition 1 at illustrative parameters ($d = 0.5$, $c = 0.5$; other values as in the replication code): outcome-detection probability required at $\beta = 0.5$ (left; values above one are infeasible) and $\beta_{\max}$ (right) against committed payment, for values of A_τ on both sides of zero.

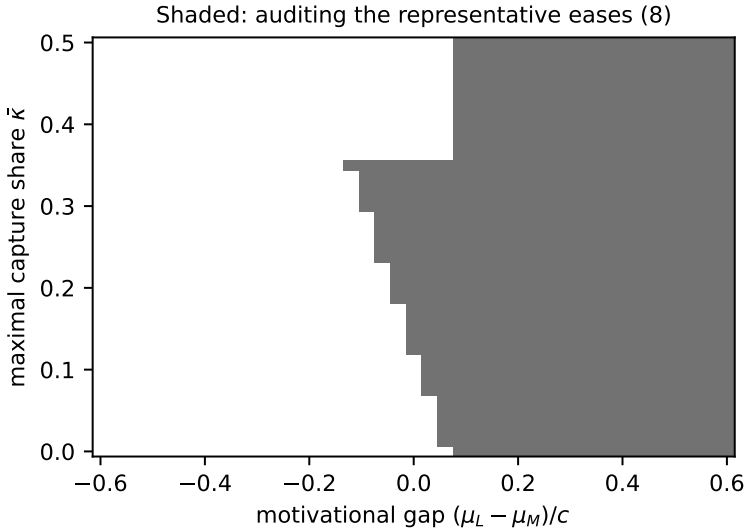


Figure 4: Proposition 4 at the first reference point with $\beta = 0.5$: shaded region in which auditing the representative eases constraint (9), by motivational gap and maximal capture share.

6 A mixed monitoring and evaluation indicator framework

The model distinguishes quantities that trigger payment, quantities that verify, quantities that govern and quantities that condition the design. A monitoring and evaluation (M&E) system for the protocol should mirror these distinctions. We propose five principles and an indicator architecture (Table 4).

Principle 1: keep payment indicators apart from learning indicators. An indicator that triggers payment inherits the integrity constraints (6) and (7): the stake riding on it must not exceed what verification can keep honest (Lemma 2), sized to the verification the funder is sure of (Lemma 4). Indicators used to learn, adapt and negotiate should not trigger payment, because the more an indicator drives decisions, the more it is exposed to distortion (Campbell, 1979). Keeping the two sets apart protects the learning system from the pressure that payment creates.

Principle 2: mix sources and methods by design. Each payment indicator should be observable through at least two independent routes, for example community monitoring and remote sensing for forest cover, or activity records and peer checks for restoration work. Triangulation raises detection, and when the second route runs through networks of peer communities rather than outside verifiers, it does so without eroding data authority (Proposition 2). Qualitative evidence, such as testimony collected in assemblies or grievance records, enters as a governance trigger rather than a payment trigger: it can open an audit, not release a payment.

Principle 3: let communities define social dimensions, and aggregate transparently. Social outcomes are multidimensional, and which dimensions matter is a community decision. A counting approach with explicit cutoffs and weights keeps aggregation transparent and decomposable, and makes explicit that the choice of dimensions, cutoffs and weights involves value judgements (Alkire and Foster, 2011). That choice therefore belongs to the negotiation of Section 7. Dimensions that are hard to verify should be paid through activity records rather than outcomes (Proposition 3).

Principle 4: give every indicator an owner, a verification route and a storage layer. For each indicator the protocol records who produces it, how it is verified, what it triggers and where it is stored. This makes the M&E system auditable and maps it directly onto the architecture of Section 8.

Principle 5: communities own and govern the capture devices. Sensors, camera traps and monitoring phones are owned, sited and maintained by the community, and their data enter community custody first; the community decides, under the data agreement of Section 7, what leaves that custody and in what form (Reyes-García et al., 2022; Schulte et al., 2025). Where certified sensing is used, readings are signed at the source and their hashes written to the record layer, which raises outcome observability (ψ) and record integrity (ω) without transferring raw data (Section 8).

In Web3 terms, this is a community-governed variant of a decentralised physical infrastructure network: participants deploy and maintain the devices, as in DePIN (The Block Research, 2026), but ownership and data-release decisions are collective and permissioned rather than individual and permissionless, because rewards paid per self-reported data point create exactly the misreporting incentive the model analyses. Networks of grassroots organisations act as non-extractive facilitators, providing training, calibration and maintenance rather than collecting data for outsiders. Paying community monitors compensates their work but does not by itself make monitoring non-extractive (Laumann et al., 2025); what does is that the monitoring system also serves the community's own decisions, which is the role of the learning indicators of Principle 1. Peer communities can certify one another's monitoring processes, in the spirit of participatory guarantee systems (May, 2019); this is the process audit of Lemma B.3, carried out by peers.

Table 4: Indicator architecture of the protocol

Family	Example indicators	Produced by	Role in the protocol	Storage
Outcome	Change in forest cover; species detections from camera traps or acoustic sensors; community-defined wellbeing index	Remote sensing; community monitors; household surveys	Triggers payment with weight β if verifiable at the funder's lower bound on detection $(y,\, D_y)$; otherwise learning	Aggregates on-chain; raw data under community custody
Activity	Patrol logs; restoration and nursery work; assemblies held	Community records, timestamped at creation	Triggers payment with weight $1-\beta$ $(z,\, q_z(\omega))$	Hashes on-chain; records off-chain
Verification	Audit coverage and detection rate; peer flag and agreement rates; share of flags resolved and time to resolution	External verifier; peer network	Sets and monitors p, ϕ and q_n; never triggers payment	Aggregate attestations on-chain
Governance and legitimacy	Auditor accepted by the assembly; device custody and data-release decisions; share of payment reaching households; participation by group; grievances filed and resolved	Assembly records; independent monitor; grievance mechanism	Tracks ℓ, κ, θ and α; can trigger an audit of representatives	Decisions on-chain; deliberations off-chain
Incentive diagnostics	External sanctions in force; programme horizon; cost of conservation work; motivation	Baseline study; structured elicitation	Estimates A_τ^0, K_T, c'_τ and μ_τ before contracting	Off-chain
Learning	Perceived fairness; trust in verifiers; knowledge exchange; unintended effects	Surveys; focus groups; assemblies	Informs renegotiation; never triggers payment	Off-chain

7 A negotiation and participation framework

The parameters of the model are not the funder's to set alone. Some require community consent to be legitimate: the degree of data authority α, the acceptance of the auditor of representatives ℓ, the benefit-sharing rule that determines capture κ, and the definition of what counts as an outcome. Others are bounded by integrity constraints that no party can negotiate away: the weight on outcomes cannot exceed the bound that verification sustains. Others are the funder's choices, such as the budget and the commitment device. The framework organises negotiation in six stages, so that each decision is taken by the party entitled to it and within the bounds the model imposes (Table 5).

Stage 1: joint diagnosis. Funder, verifier and community exchange the information needed to locate the setting: the external sanctions that apply, the programme horizon, the cost of the conservation work and the risk of capture (Section 9). The community participates through its own governance bodies, not through a single representative. In the Colombian case of Section 1, the developer relied on a mandate contract signed with one legal representative, and the Court found that genuine free, prior and informed consent had not been obtained (Corte Constitucional de Colombia, 2024).

Stage 2: what counts. The community proposes outcome definitions, social dimensions and weights, and the activity records it will keep; the verifier states how each can be verified and at what detection; the funder states what it is willing to pay for. Indicators that cannot be verified at the required detection move to activity payment or to the learning set (Section 6).

Stage 3: data authority. The community decides which territorial data remain under its custody and on what terms outsiders may access them. The verifier states the detection this allows. If it is too low, the parties add a network of peer communities with audited checks (Proposition 2; Lemmas B.1–B.3) rather than overriding data authority.

Stage 4: accountability. The community's governance bodies accept or reject the proposed auditor of representatives, agree the benefit-sharing rule and establish a grievance mechanism. Where the state has a supervisory role, the auditor coordinates with it (Proposition 4).

Stage 5: contract terms. Within the bounds set by Stages 1–4, the funder proposes the weight on outcomes, verification intensity, number of peers, sanctions, horizon and commitment device, and the community consents. The model supplies the bounds: $\beta \leq \beta_{\max}(\underline{D})$ (Lemma 4), the number of peers required by Lemma B.1, and stronger verification where commitment is credible and $A_\tau^0 > 0$ (Proposition 1). The output is a term sheet signed by the three parties.

Stage 6: review. At fixed intervals, learning indicators and verification statistics inform renegotiation of the terms, under the same consent rules.

Table 5: Negotiation matrix

Parameter	Proposed by	Consent required from	Bound from the model	Recorded in
Outcome definitions; social dimensions, cutoffs and weights	Community	Funder and verifier, on verifiability	Payment indicators verifiable at $\underline{D}$	Term sheet; hash on-chain
Data authority α	Community	Community decision	Add audited peer network if detection is too low (Prop. 2)	Data agreement; hash on-chain
Auditor of representatives; benefit-sharing rule	Funder or verifier; community	Community governance bodies	Prop. 4; Cor. 1	Assembly decision on-chain
Weight on outcomes β	Funder	Community	$\beta \leq \beta_{\max}(\underline{D})$ (Lemma 4)	Smart contract
Verification p; number of peers k	Verifier	Funder and community	Lemmas B.1–B.3	Smart contract; attestations
Sanctions; horizon T; commitment δ	Funder	Community	Prop. 1; Table 7	Smart contract; escrow

8 Coupling the protocol to a blockchain

The framework is chain-agnostic: it specifies functions, not a particular network. Figure 5 shows five on-chain layers and what stays off-chain.

Rules and payment layer. A smart contract encodes the payment rule (1), holds the budget in escrow and releases payments against verified attestations; withholding and exclusion after detected misreporting are executed by rule. This layer implements commitment ($\delta = 0$). By Proposition 1, it must be paired with stronger verification where $A_\tau^0 > 0$ or the programme is short.

Record layer. Cryptographic hashes of activity records and of the documentation of peer-triggered checks are written when the records are created, while the records themselves stay off-chain. This layer implements record integrity (ω): it raises the detection of fabricated records and makes the audit of network checks feasible (Lemma B.3).

Attestation layer. Verifiers and peers submit signed attestations about outcomes and checks. Peer identities are protected by pseudonymous keys and by random, non-reciprocal assignment, which removes the gain from reciprocal leniency (Lemma B.2); the source of randomness must itself be trustworthy. Verified claims can be represented as impact certificates that admit several evaluations (Gitcoin, 2026).

Disclosure layer. Rules, payment flows, aggregated verified outcomes and verification statistics are public. Raw territorial data remain in community-controlled storage, with access governed by the data agreement of Stage 3 (α), consistent with the CARE principles (Carroll et al., 2020).

Governance layer. Community decisions on which the protocol depends, such as acceptance of the auditor and renegotiation, are recorded on-chain as decisions of the community's own bodies, not as token-weighted votes, whose concentration is well documented (Feichtinger et al., 2024). Where coercion or bribery of voters is a risk, which ledger-based voting can make realistic (Fábrega et al., 2025), individual votes should be kept private.

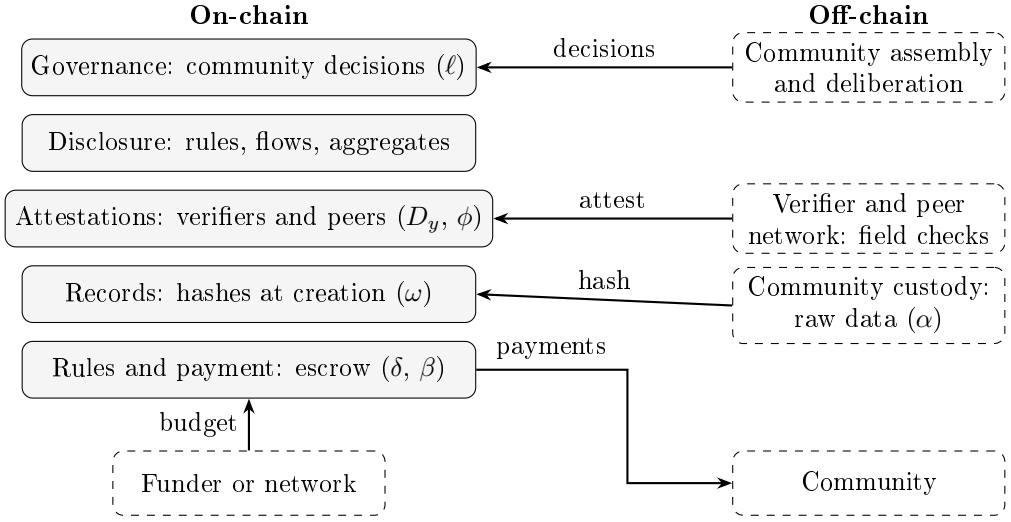


Figure 5: Chain-agnostic architecture. Ledgers provide commitment and record integrity and host the public disclosure layer; they do not observe the territory, which remains under community custody.

Trust assumptions and threat model. Putting more of the protocol on-chain does not remove every point of trust. Four remain and are the points to audit: the external verifier's random sample and its audit of network checks; the source of randomness for peer assignment; the legitimacy of the community bodies whose decisions are recorded on-chain, since who counts as the community is decided off-chain in the negotiation of Section 7; and the legal enforceability of sanctions outside the programme (S_{ext}). Table 6 lists the attack vectors the design addresses and where each is handled.

What the chain cannot do follows from the model. It does not make territorial outcomes observable, it does not supply legitimacy, and its commitment function does not by itself improve integrity: escrow accounts and trusts with legal enforcement provide commitment too, and every result about commitment applies to them. The distinctive contribution of a ledger is the joint provision of commitment and record integrity, which are complementary where $A^0_\tau > 0$ (Proposition 1b), together with a public disclosure layer that all parties can audit.

Table 6: Threat model: attack vectors and mitigations

Attack vector	Mitigation in the protocol	Where
Outcome inflation by the reporting community	Payout on outcomes capped by the verification the funder is sure of; hybrid payouts; peer challenges	Lemmas 2–4; Props. 1, 3
Fabricated activity records	Records hashed at creation; record audits	Constraint (7); record layer
Verifier capture (verifier paid by the party it audits)	Verifier selected and paid under the term sheet, not by the project developer; audit of verification processes	Section 7; Lemma B.3
Peer collusion and reciprocal leniency	Anonymous, random, non-reciprocal assignment	Lemma B.2
Collusion in network checks	External audit of documented checks	Lemma B.3
Sybil peers (fictitious communities or verifiers)	Admission of peers through the governance layer and community identity attestations	Trust assumption; not modelled
Elite or governance capture	Audits of representatives through auditors the community accepts; benefit-sharing rule on-chain	Prop. 4; Cor. 1
Vote buying and coercion	Private voting; decisions by community bodies, not token-weighted	Governance layer
Data extraction	Community custody; access by agreement; only aggregates on-chain	Principle 5; α
Sensor spoofing	Signed-at-source devices; triangulation with an independent route	Principle 2; ψ

9 From framework to field

This section turns the framework into a pilot. It is written for a setting such as the Colombian reference case and indicates how each decision simplifies where the corresponding complexity is absent. It is a design to be tested, not a validated standard.

9.1 Diagnosing the setting

Diagnosis is not only quantitative. It combines an anthropological and social study of the territory, covering its governance bodies, livelihoods, knowledge systems, internal conflicts and security conditions, with an analysis of the macroeconomic and social context, including markets for the alternatives to conservation, relevant public policy and labour conditions. These studies determine the values of the model's parameters, which this document deliberately leaves open: payment levels, remuneration, horizons and timelines are set per project and per territory.

Four quantities determine which design can work, and each can be approximated before a pilot begins: the sign of A_τ^0; outcome informativeness $\Delta\pi$, that is, how strongly the outcome to be paid for responds to the community's effort rather than to exogenous drivers; the risk of capture $\bar{\kappa}$ and the accountability of representatives to members; and the community's preferred degree of data authority α, which should be agreed rather than assumed. The condition in Proposition 1 compares external sanctions with $K_T c'_\tau$, the present value of the net cost of the conservation work a community would stop doing if excluded, measured in periods of that cost. Table 7 reports K_T for annual rounds at three discount rates and three programme horizons. In a ten-year programme at a 10 percent discount rate, for example, commitment complements verification only if external sanctions exceed about six years of the net cost of conservation work. Two implications follow without calibration. Where work is costly and intrinsic motivation modest, the threshold is demanding, so long programmes are likely to lie in the region where credible payment disciplines through rents. Short programmes, impatience and high intrinsic motivation all lower $K_T c'_\tau$ and push toward complementarity; pilots and short contracts, common when protocols are first tested, are therefore where committed payment most needs strong verification. Locating actual programmes requires data on payments, costs and sanctions that we do not have: the Delphi study elicits the sign, and calibration is left to the companion empirical work.

Table 7: Discounted number of future rounds lost on exclusion, $K_T = d(1-d^T)/(1-d)$ with $d = 1/(1+r)$

Discount rate r	$T = 5$	$T = 10$	$T = \infty$
5%	4.33	7.72	20.00
10%	3.79	6.14	10.00
20%	2.99	4.19	5.00

9.2 Designing the pilot

The pilot follows six steps. First, diagnose the setting as above. Second, agree with the community what counts as an outcome, which social dimensions matter and which activity records will be kept, and assign every indicator its role (Section 6). Third, agree data authority and, where outside observability is then too low, set up a network of peer communities with audited checks (Section 7; Lemmas B.1–B.3). Fourth, agree accountability: an auditor of representatives accepted by the community's governance bodies, a benefit-sharing rule and a grievance mechanism (Proposition 4). Fifth, set contract terms within the model's bounds, with the weight on outcomes at the bound implied by the verification the funder is sure of (Lemma 4), and deploy the architecture of Section 8, pairing credible payment with record integrity and stronger verification where $A_\tau^0 > 0$ or the programme is short. Sixth, evaluate, as described below. Table 8 summarises the decisions and how they simplify in less demanding settings.

Security risks, which are central in the reference case, enter the model only indirectly: they raise the cost of field verification and audits, lowering the detection probabilities a funder can secure, and they make anonymity of peer verifiers harder to guarantee and more important.

Table 8: Design decisions for a pilot

Decision	Model result	Measure before the pilot	Where conditions are simpler
Payment object (β)	Prop. 3; Lemmas 3–4	Outcome informativeness $\Delta\pi$; lower bound on detection	Where outcomes are observable externally at low cost, more weight on outcomes is safe
Verification (p, ϕ, α)	Prop. 2; Lemmas B.1–B.3	Preferred data authority; peer accuracy; audit capacity	Where communities accept outside access to territorial data, direct verification can replace peer networks
Accountability (τ, ℓ)	Prop. 4; Cor. 1	Capture risk; acceptance of the auditor	Where representatives are directly accountable and capture risk is low, member-level compliance checks suffice
Commitment (δ, ω)	Prop. 1; Table 7	Sign of A_τ^0; programme horizon	In long programmes that are the community's main source of net surplus, commitment partly substitutes for verification

9.3 Evaluation

Following the design-science sequence, the numerical illustration of Section 5 demonstrates the artefact and the Delphi study evaluates it (Peffers et al., 2007). We use the Delphi technique, an iterative and anonymous method for gathering and evaluating expert knowledge whose main advantage over other group techniques is that it reduces social pressure among respondents (Mukherjee et al., 2015). The protocol accompanies this paper (Supplementary Appendix D) and will be registered on OSF before recruitment.

The panel is structured by five stakeholder groups (researchers; representatives of grassroots

organisations, participating as remunerated experts; current and former regulators in a personal capacity; practitioners of certification standards; and ledger developers), because the protocol's legitimacy depends on whether each group finds its assumptions plausible, not on an aggregate average. Over three rounds, panelists rate the plausibility of the maintained assumptions, rate protocol requirements and flag incompatible pairs, and judge the quantities on which the main result turns. The sign of the benchmark A_τ^0 is elicited twice as a convergent-validity check: directly, by asking whether guaranteed payment would change a community's incentive to report honestly, and comparatively, by weighing the outside consequences of detected misreporting against the savings from no longer doing the programme's conservation work. Motivational responses to leader audits, member audits and collective payment are elicited with the four-step format, which across three studies was more likely to reduce overconfidence than a three-point procedure (Speirs-Bridge et al., 2010).

At panel level, consensus requires at least 70 percent of respondents in the top two categories and an interquartile range of at most one point; within groups, it requires minimum counts fixed in the protocol in at least four of the five groups. Results are reported by group, dissent is described, and assumptions without consensus in two or more groups are treated as principal limitations. Judgements enter the numerical exploration through mapping rules fixed before collection: S_{ext} is sampled relative to Kc'_τ so that A_τ^0 takes only the signs contained in at least one group's interquartile range on the comparative item; θ and S_R, which the items do not address, are sampled over their Table 3 ranges and reported separately; and the sign region of $\mu_L - \mu_M$ is restricted in the same way.

Grassroots organisations take part through an independent network of such organisations, whose role in the study is facilitative and non-extractive: they convene participants, adapt the instrument to local languages and forms of deliberation, and return results to the communities. Because the author's firm has a commercial interest in verification protocols, facilitation, aggregation and coding are carried out by facilitators without a commercial stake in that firm and by a second coder, and the author sees only anonymised results. The study will undergo independent ethics review and follow the consent protocols of participating organisations and communities before recruitment. Consensus is not validity, and panel selection shapes results; we therefore publish recruitment criteria and the realised composition.

The model yields four predictions for a framed field experiment with community groups, to be pre-registered separately. Holding verification constant, making payment certain rather than probabilistic should increase inflation of self-reported outcomes when external sanctions are salient and decrease it when exclusion from future rounds is the main deterrent (Proposition 1); crossing payment certainty with the salience of external sanctions tests the sign condition directly. At equal expected budget, hybrid payment should reduce inflation relative to pure payment by results (Proposition 3). Audits of representatives by an auditor the group has accepted should reduce capture and raise member effort relative to audits of members, whereas an imposed auditor should weaken or reverse this effect (Proposition 4). And multi-task verification between groups should reduce infla-

tion even when outside observability is low (Proposition 2). A design crossing payment certainty with external-sanction salience and with payment object, with an additional auditor-legitimacy arm, tests the first three directly.

9.4 Implementation phases

A field implementation can proceed in five phases, each with a verifiable deliverable. Phase 0, diagnosis, produces the anthropological, social and contextual studies, the diagnostic quantities above and a map of the community's governance bodies. Phase 1, co-design and negotiation, produces the term sheet of Section 7, signed by funder, community and verifier. Phase 2, baseline and set-up, produces baseline values of payment and learning indicators, the peer network and the deployed contracts, with record and data agreements in place. Phase 3, operation, releases payment rounds against verified attestations, with audits and peer checks at the agreed intensities. Phase 4, evaluation and renegotiation, produces the Delphi results, the pilot's verification statistics and learning indicators, and a revised term sheet. Capital is released only against verified results, and each phase produces evidence that refines the parameters of the next.

9.5 Compliance mapping

The protocol is not a legal instrument, but its design elements map onto requirements that projects must meet. Table 9 relates them to Colombian norms, as an example of a demanding national framework, and to international references. Each project requires its own legal review in its jurisdiction.

Table 9: Compliance mapping: Colombian example and international references

Requirement	Colombian example	International reference	Protocol element
Consent of affected communities through their own institutions	Ruling T-248 of 2024: genuine free, prior and informed consent and an ethnic approach in carbon projects (Corte Constitucional de Colombia, 2024)	Standards that require free, prior and informed consent, such as the CCB Standards (Verra)	Negotiation Stages 1–4; consent recorded from community governance bodies
Protection of personal data of monitors and households	Law 1581 of 2012, general regime for personal data protection (Congreso de la República de Colombia, 2012)	Data protection law of each jurisdiction	Data minimisation; community custody; access agreements
Collective and Indigenous data governance	Self-government rights protected in ruling T-248 (Corte Constitucional de Colombia, 2024)	CARE principles (Carroll et al., 2020)	Data authority α; device custody; disclosure layer
Environmental information, participation and protection of defenders	Escazú Agreement, approved by Law 2273 of 2022; Colombia a party since December 2024 (Ministerio de Ambiente y Desarrollo Sostenible de Colombia, 2024)	Regional agreement for Latin America and the Caribbean (Ministerio de Ambiente y Desarrollo Sostenible de Colombia, 2024)	Public disclosure layer; grievance mechanism; protection of monitors' identities
Benefit sharing and independent verification	Ruling T-248 on state supervision of the voluntary carbon market (Corte Constitucional de Colombia, 2024)	Plan Vivo: at least 60 percent of income to communities, third-party verification (Plan Vivo Foundation)	Benefit-sharing rule; verifier role; audits of representatives

10 Limitations and scope

The model is deliberately minimal. The representative's regard for members is exogenous; endogenising it through local accountability institutions would link Corollary 1 and Proposition 4. The funder takes the budget and verification intensity as given; pricing verification and audits would turn the design rules of Section 9 into an optimal design problem. The protocol is stationary with permanent exclusion and no renegotiation; non-stationary payments, temporary exclusion or reputational dynamics across funders would change the fixed part of the sanction and could move settings across the condition in Proposition 1. The analysis of peer verification abstracts from coalitions that make side payments and from signals correlated across neighbouring territories, and it assumes that the checks peer flags trigger can be externally audited as documented processes. Detection is linear in data authority and effort is binary, although the planned exploration relaxes the former. There is no market for credits: if buyers pay premiums for verified credits, verification investments become partly self-financing, which our single-funder setting cannot capture. The parameter ranges are stylised, bounded by orders of magnitude from different settings and informed by the Delphi study rather than calibrated, and the Delphi study cannot replace deliberation with serving regulators.

The framework takes results-based payment as given and asks how to make it truthful and fair; it does not establish that such payment is always the right instrument. Where communities would conserve without payment, the protocol verifies transfers rather than incentives, and its justification is benefit sharing rather than additionality (Section 4). The framework also takes the outcome as defined: where communities define what counts, truthful reporting is relative to that agreed definition, which is why Section 7 places the definition in the negotiation.

11 Conclusion

Web3 infrastructure can make payments for nature-based solutions credible and their records tamper-evident, but it cannot make territorial outcomes observable or verification legitimate. Community verification must therefore be designed, and this paper has provided the elements of that design. Without future rents, verification capacity, not budget, limits how much payment can depend on self-reported outcomes. Credible on-chain payment complements verification when a community's sanctions outside the programme exceed the discounted cost of the conservation work it would stop doing if excluded, and substitutes for it otherwise; elite capture pushes toward complementarity. Community data authority and verifiability can be reconciled through audited networks of peer communities, and accountability works best when it targets representatives through auditors that communities accept.

On this basis, the framework offers funders and networks a mixed indicator system that keeps payment and learning apart, a negotiation process that gives each party the decisions it is entitled to, and an architecture that puts rules, payments and proofs on-chain while leaving territorial data with communities. Colombia's combination of collective rights, constitutional oversight and security risks makes it a demanding test; settings with fewer of these layers should find the requirements

easier to meet. The next step is a field pilot, which will test the model's predictions and refine its parameters.

Declarations

Status. This whitepaper sets out a methodological framework for a field pilot. It specifies no payment levels, prices, remuneration or timelines, which are set per project after the diagnosis of Section 9. It has not been peer reviewed and will be revised with evidence from the Delphi study and the pilot.

Author contributions. Conceptualisation, methodology, formal analysis, software and writing: R. S. Villalba Arango.

Competing interests. The author is the founder of Verified Impact Protocol (VIP), which intends to offer implementation services based on this framework.

Funding. This work received no funding.

Data and code availability. The model code, unit tests, figure scripts and the Delphi protocol are provided as supplementary files with this paper. No data from human participants were collected for this paper.

Use of AI tools. An AI assistant was used to search for and verify literature, draft text and check derivations and code. All analytical choices, interpretations and the final text are the author's responsibility.

References

Sabina Alkire and James Foster. Counting and multidimensional poverty measurement. *Journal of Public Economics*, 95(7–8):476–487, 2011. doi: 10.1016/j.jpubeco.2010.11.006.

Gary S. Becker. Crime and punishment: An economic approach. *Journal of Political Economy*, 76 (2):169–217, 1968.

Lucas Belenky, Keisuke Iyadomi, Susan David Carevic, and Harikumar Gadde. Digital monitoring, reporting, and verification systems and their application in future carbon markets. Technical report, World Bank, 2022.

Martina Björkman and Jakob Svensson. Power to the people: Evidence from a randomized field experiment on community-based monitoring in Uganda. *Quarterly Journal of Economics*, 124(2): 735–769, 2009. doi: 10.1162/qjec.2009.124.2.735.

Vitalik Buterin, Zoë Hitzig, and E. Glen Weyl. A flexible design for funding public goods. *Management Science*, 65(11):5171–5187, 2019. doi: 10.1287/mnsc.2019.3337.

Giulio Caldarelli. Understanding the blockchain oracle problem: A call for action. *Information*, 11 (11):509, 2020. doi: 10.3390/info11110509.

Donald T. Campbell. Assessing the impact of planned social change. *Evaluation and Program Planning*, 2(1):67–90, 1979. doi: 10.1016/0149-7189(79)90048-X. Open reprint: Journal of MultiDisciplinary Evaluation 7(15), 2011, doi:10.56645/jmde.v7i15.297.

Stephanie Russo Carroll, Ibrahim Garba, Oscar L. Figueroa-Rodríguez, Jarita Holbrook, Raymond Lovett, Simeon Materechera, Mark Parsons, Kay Raseroka, Desi Rodriguez-Lonebear, Robyn Rowe, Rodrigo Sara, Jennifer D. Walker, Jane Anderson, and Maui Hudson. The CARE principles for indigenous data governance. *Data Science Journal*, 19:43, 2020. doi: 10.5334/dsj-2020-043.

Congreso de la República de Colombia. Ley estatutaria 1581 de 2012, por la cual se dictan disposiciones generales para la protección de datos personales, 2012. Diario Oficial 48.587, 18 October 2012.

Corte Constitucional de Colombia. Sentencia T-248 de 2024. `https://www.corteconstitucional.gov.co/relatoria/2024/t-248-24.htm`, 2024. Sala Segunda de Revisión, M.P. Juan Carlos Cortés González, expediente T-9.312.858, 25 June 2024.

Finn Danielsen, Per M. Jensen, Neil D. Burgess, et al. A multicountry assessment of tropical resource monitoring by local communities. *BioScience*, 64(3):236–251, 2014. doi: 10.1093/biosci/biu001.

Anirban Dasgupta and Arpita Ghosh. Crowdsourced judgement elicitation with endogenous proficiency. In *Proceedings of the 22nd International Conference on World Wide Web (WWW '13)*, pages 319–330, 2013.

Andrés Fábrega, Amy Zhao, Jay Yu, James Austgen, Sarah Allen, Kushal Babel, Mahimna Kelkar, and Ari Juels. Voting-bloc entropy: A new metric for DAO decentralization. In *Proceedings of the 34th USENIX Security Symposium*, 2025.

Rainer Feichtinger, Robin Fritsch, Yann Vonlanthen, and Roger Wattenhofer. The hidden shortcomings of (D)AOs – an empirical study of on-chain governance. In *Financial Cryptography and Data Security. FC 2023 International Workshops*, volume 13953 of *Lecture Notes in Computer Science*, pages 165–185. Springer, 2024. doi: 10.1007/978-3-031-48806-1.

Gitcoin. Impact certificates (hypercerts). `https://gitcoin.co/mechanisms/impact-certificates-hypercerts`, 2026.

Paula Helm, Amalia de Götzen, Luca Cernuzzi, Alethia Hume, Shyam Diwakar, Salvador Ruiz Correa, and Daniel Gatica-Perez. Diversity and neocolonialism in big data research: Avoiding extractivism while struggling with paternalism. *Big Data & Society*, 2023. doi: 10.1177/20539517231206802.

Bengt Holmström and Paul Milgrom. Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics, and Organization*, 7:24–52, 1991.

Seema Jayachandran, Joost de Laat, Eric F. Lambin, Charlotte Y. Stanton, Robin Audy, and Nancy E. Thomas. Cash for carbon: A randomized trial of payments for ecosystem services to reduce deforestation. *Science*, 357(6348):267–273, 2017. doi: 10.1126/science.aan0568.

Goytom Abraha Kahsay and Erwin Bulte. Internal versus top-down monitoring in community resource management: Experimental evidence from Ethiopia. *Journal of Economic Behavior & Organization*, 189:111–131, 2021. doi: 10.1016/j.jebo.2021.06.030.

Goytom Abraha Kahsay, Erwin Bulte, Francisco Alpizar, Lars G. Hansen, and Haileselassie Medhin. Leadership accountability in community-based forest management: Experimental evidence in support of governmental oversight. *Ecology and Society*, 28(4):20, 2023. doi: 10.5751/ES-14469-280420.

Katie May Laumann, Nicholas M. Hoad, Lauren Alvaro, Shahrzad Lili Badri, Noirin Burke, Annie Carew, Guilherme N. Corte, Aldo Croquer, Yasmina Shah Esmaeili, Martha Farrell, Naoko Kouchi, Juhyung Lee, Masahiro Nakaoka, Lina Mtwana Nordlund, Rita I. Sellares-Blasco, Ed Sheldon, Maria F. Villalpando, and Jonathan S. Lefcheck. Closing the parachute and opening the umbrella: Strategies for inclusivity and representation in producing impactful coastal ecosystem research. *Cambridge Prisms: Coastal Futures*, 2025. doi: 10.1017/cft.2025.10004.

Tim Mackey, Alec J. Calac, Tiana J. McMann, Ken Miyachi, Samm Hurst, Keolu Fox, Burt Dillabaugh, Kelle Dhein, Guthrie Ducheneaux, and Joe Yracheta. Development of a blockchain-based platform to enable indigenous data sovereignty and shared research participation with indigenous communities: Technology prototyping and community engagement study. *Journal of Medical Internet Research*, 28:e90247, 2026. doi: 10.2196/90247.

Tim K. Mackey, Alec J. Calac, B. S. Chenna Keshava, Joseph Yracheta, Krystal S. Tsosie, and Keolu Fox. Establishing a blockchain-enabled indigenous data sovereignty framework for genomic data. *Cell*, 185(15):2626–2631, 2022. doi: 10.1016/j.cell.2022.06.030.

Chris May. PGS guidelines: How to develop and manage participatory guarantee systems for organic agriculture. IFOAM – Organics International; FAO, 2019. Quotes the IFOAM (2008) definition of participatory guarantee systems.

Ministerio de Ambiente y Desarrollo Sostenible de Colombia. Acuerdo de Escazú. `https://www.minambiente.gov.co/acuerdo-de-escazu/`, 2024. Approved by Law 2273 of 2022; declared constitutional in ruling C-359 of 2024; Colombia a State Party since 24 December 2024.

Lina Moros, María Alejandra Vélez, and Esteve Corbera. Payments for ecosystem services and motivational crowding in Colombia's Amazon Piedmont. *Ecological Economics*, 156:468–488, 2019. doi: 10.1016/j.ecolecon.2017.11.032.

Nibedita Mukherjee, Jean Hugé, William J. Sutherland, Jeffrey McNeill, Maarten Van Opstal, Farid Dahdouh-Guebas, and Nico Koedam. The Delphi technique in ecology and biological conservation:

Applications and guidelines. *Methods in Ecology and Evolution*, 6(9):1097–1109, 2015. doi: 10.1111/2041-210X.12387.

Benjamin A. Olken. Monitoring corruption: Evidence from a field experiment in Indonesia. *Journal of Political Economy*, 115(2):200–249, 2007.

Optimism Collective. RetroPGF 3: Round design. `https://gov.optimism.io/t/retropgf-3-round-design/6802`, 2023.

Payne Institute for Public Policy. Voluntary carbon markets' uneasy start with blockchain. `https://payneinstitute.mines.edu/voluntary-carbon-markets-uneasy-start-with-blockchain/`, 2022. Summarises a Carbon Direct analysis of Verra registry retirements, 4Q2021–1Q2022.

Ken Peffers, Tuure Tuunanen, Marcus A. Rothenberger, and Samir Chatterjee. A design science research methodology for information systems research. *Journal of Management Information Systems*, 24(3):45–77, 2007. doi: 10.2753/MIS0742-1222240302.

Plan Vivo Foundation. PV Climate: About. `https://www.planvivo.org/pv-climate-about`. Institutional page; cite the PV Climate Standard (v5.0) document in the final version.

Victoria Reyes-García, Adrien Tofighi-Niaki, Beau J. Austin, Petra Benyei, Finn Danielsen, Álvaro Fernández-Llamazares, Aditi Sharma, Ramin Soleymani-Fard, and Maria Tengö. Data sovereignty in community-based environmental monitoring: Toward equitable environmental data governance. *BioScience*, 72(8):714–717, 2022. doi: 10.1093/biosci/biac048.

Julian Rode, Erik Gómez-Baggethun, and Torsten Krause. Motivation crowding by economic incentives in conservation policy: A review of the empirical evidence. *Ecological Economics*, 117:270–282, 2015. doi: 10.1016/j.ecolecon.2014.11.019.

Kayla Schulte, Andrew Grieve, Benjamin Barratt, Timothy Baker, Hima Coonjobeeharry, and Mohammed Mead. Advancing participatory sensing and knowledge production methods for city air quality governance: Applying the Breathe London community programme model. *Environmental Science & Policy*, 170:104092, 2025. doi: 10.1016/j.envsci.2025.104092.

Carl Shapiro and Joseph E. Stiglitz. Equilibrium unemployment as a worker discipline device. *American Economic Review*, 74(3):433–444, 1984.

Victor Shnayder, Arpit Agarwal, Rafael Frongillo, and David C. Parkes. Informed truthfulness in multi-task peer prediction. *arXiv preprint*, (1603.03151), 2016.

Margaret Skutsch, Michael K. McCall, and Alejandro P. Larrazabal. Balancing views on community monitoring: The case of REDD+. *Biodiversity and Conservation*, 23(1):233–236, 2014. doi: 10.1007/s10531-013-0594-1.

Tara Slough, Jacob Kopas, and Johannes Urpelainen. Satellite-based deforestation alerts with training and incentives for patrolling facilitate community monitoring in the Peruvian Amazon. *Proceedings of the National Academy of Sciences*, 118(29):e2015171118, 2021. doi: 10.1073/pnas.2015171118.

Birte Snilstveit, Jennifer Stevenson, Laurenz Langer, Natalie Tannous, Zafeer Ravat, Promise Nduku, Joshua Polanin, Ian Shemilt, John Eyers, and Paul J. Ferraro. Incentives for climate mitigation in the land use sector: The effects of payment for environmental services on environmental and socioeconomic outcomes in low- and middle-income countries: A mixed-methods systematic review. *Campbell Systematic Reviews*, 15(3):e1045, 2019. doi: 10.1002/cl2.1045. Author list as in the published article; the 3ie report version lists two authors under other surname forms.

Derek Sorensen. Tokenized carbon credits. *Ledger*, 8:76–91, 2023. doi: 10.5195/ledger.2023.294.

Andrew Speirs-Bridge, Fiona Fidler, Marissa McBride, Louisa Flander, Geoff Cumming, and Mark Burgman. Reducing overconfidence in the interval judgments of experts. *Risk Analysis*, 30(3): 512–523, 2010. doi: 10.1111/j.1539-6924.2009.01337.x.

The Block Research. The DePIN report 2025. `https://www.theblock.co/post/360958/state-of-depin-2025`, 2026.

Robert M. Townsend. Optimal contracts and competitive markets with costly state verification. *Journal of Economic Theory*, 21(2):265–293, 1979. doi: 10.1016/0022-0531(79)90031-0.

Verra. Climate, community & biodiversity standards program. `https://verra.org/programs/ccbs/`.

Thales A. P. West, Jan Börner, Erin O. Sills, and Andreas Kontoleon. Overstated carbon emission reductions from voluntary REDD+ projects in the Brazilian Amazon. *Proceedings of the National Academy of Sciences*, 117(39):24188–24194, 2020. doi: 10.1073/pnas.2004334117.

World Bank Financial Intermediary Funds. Tropical forest forever facility. `https://fiftrustee.worldbank.org/en/about/unit/pim/fiftrustee/fund-detail/tfff`, 2026. Trustee page; launched at COP30 (Belém, 2025).