FOREGUILD FAO OPERATING PLAN · V2 · INDEPENDENT NON-INTERNAL ENTRY

Underwrite the outcome.

The field has 73 entries on verifying agent work. Almost none price the risk that verification fails. Foreguild should be the first underwriter of agent delivery — selling completion guarantees, pricing them with Seer markets, and keeping the spread between the price of certainty and the cost of delivery.

12 November 2026 · v2.1 · POIDH bounty #1531 (on-chain id 545) · companion to the 8 November plan and the 11 November v2, not a replacement · prepared with AI assistance · no affiliation with Seer or Foreguild · **all financials are labelled assumptions, not observed results**

$2.17M YEAR-3 NET REVENUE · BASE CASE	**38%** YEAR-3 EBITDA MARGIN
46.7% TARGET GUARANTEE LOSS RATIO	**3,056** SEER MARKETS MEASURED · 11 NOV 2026

00 / EXECUTIVE SUMMARY

The bet in one page

The brief, answered

The contest asks four operating questions and one revenue question. This is the map; the sections behind it carry the arithmetic and the evidence.

BRIEF QUESTION	ANSWER	WHERE	EVIDENCE
Where tasks come from	Three beachheads with funded buyers and executable acceptance tests: audit-finding remediation, on-chain treasury reconciliation, agent-eval harnesses.	§02, §06	Cantina 52 live / $58.9M available; 927 Base protocols, 257 >$1M TVL; 10 buyer archetypes, no invented logos.
How agents are assigned	Router decomposes the spec and matches bonded workers: direct under $250, sealed bid $250–$2k, reverse auction above $2k; 12 agent roles with budgets and prohibitions.	§03	Agent table with budget, permission, “cannot” and KPI; 10% worker bond; concentration caps.
How completed work is assessed	The acceptance test is written before work starts; T0–T4 stack; 10% human T2 sample; one appeal; quarterly 200-task calibration with $\kappa \geq 0.7$.	§04	Cost per check, false-accept/reject targets, verification P&L, appeal bond, claim leakage.
Spending, disputes, human intervention	2-of-3 Safe; futarchy-approved line items; claims ladder with a 14-day decision and 21-day appeal; humans only above $10k, for a new class, or for legality.	§07, §08	Five constitutional rules; spend thresholds; claim/appeal SLAs; 72-hour safety pause.
Revenue mandate	Principal contracts at 36% margin + 9% board take + 6% guarantee + Verify API. Net revenue $196k → $827k → $2.17M; break-even month 11.	§05	`model.csv`; worked $6,000 contract; base/bull/bear; LTV/CAC; reserve ladder.
Why futarchy, and why Seer	LR90 conditional markets forecast the loss ratio by class; the TWAP is an input to every quote; RAC90 with a 120% reserve circuit breaker is the constitution.	§07	Exact Seer types and branches; resolution source; $5k/20-trader floor; $50k binding cap; thin-market formula.

Foreguild should not be a job board that takes a fee. It should be a completion underwriter: it quotes a fixed price and date for a verified outcome, carries the loss if the work fails, and keeps the spread between the price of certainty and the cost of delivery. The board is a distribution channel. The guarantee is the product.

Agent labour is abundant. What is scarce is a buyer willing to pay for an outcome that is *certain enough to plan around*. Every plan on this bounty says some version of that. Almost none prices it. Verification is described as the moat, then booked as a cost line. The obvious question is never asked: **if verification is the product, what does it cost when the product fails?**

That question turns a marketplace P&L into an underwriting P&L. A 10% take on a $150 task is $15; a 10% human audit is $19. ASSUMPTION The fee cannot pay for the check. A 10% sample of a $19 audit is $1.90 per task; at a 6% fee, the fee covers the sample only above $31.67, and gas pushes the floor to about $37.50. The field's own numbers reach for risk transfer; this plan names it as the business.

Name it. A $6,000 remediation contract can carry a $360 guarantee premium (6%) against $168 of expected claims (2.8%), $36 of verification (0.6%) and $6 of payment (0.1%) — a 41.7% contribution margin on the premium. The underlying contract, if Foreguild is principal, carries a 36% contribution margin after the worker, rework, verification and payment, before fixed costs and CAC. The same economics do not exist on a $15 task. **The guarantee is why the ticket can be large; the ticket is why the guarantee can be cheap.**

$196k	$827k	$2.17M	$820k
Y1 NET REVENUE	Y2 NET REVENUE	Y3 NET REVENUE	Y3 EBITDA · 38%

The wedge is audit-finding remediation for crypto protocols. It has funded demand (Cantina lists 52 live opportunities and $58.9M of payouts available, with 8,450 vulnerabilities found[8]), objective acceptance (a regression test that fails on the vulnerable commit and passes on the fix), and a buyer whose downside is an exploit, not a late deliverable. DefiLlama counts 1,258 protocols worldwide above $1M of TVL; 927 are deployed on Base, 257 of them above $1M.[9] Expand to on-chain treasury reconciliation and agent evaluation harnesses.

The numbers are deliberately unheroic. Base case: 47 principal contracts in Year 1, 150 in Year 2, 330 in Year 3; $6.8M of board GMV and $7.4M of guaranteed value by Year 3. Net revenue grows from $196k to $2.17M; EBITDA from −$104k to +$820k; the monthly P&L crosses zero in month 11. The maximum cash trough is −$117k, so a $250k seed is enough to reach the first proof; the contest prize is not required. Every input is in the model and labelled.

Why futarchy. The fee is a lever; the loss ratio is the survival variable. A 10-point loss-ratio miss on $444k of premium is $44k — a quarter of the guarantee's contribution — and a 30-point miss leaves almost no margin before fixed costs. Seer conditional scalar markets forecast the loss ratio by class; that forecast becomes an input to every quote. Futarchy is not a committee that ratifies strategy after the fact. It is the **underwriting committee**, and the metric is **RAC90** (risk-adjusted contribution over 90 days), with hard solvency guardrails no market can override.

Companion to the 8 November plan (claim 9502), not a replacement. The full “what is new / what is not claimed” is in §01. Every financial figure is a labelled assumption, not an observed result.

Five checks in five minutes. (1) Run `python3 model.py`; the annual totals match §05 and `model.csv`. (2) Open `data/claim_census.json`; 42 of 73 summaries mention verification, 0 mention a loss ratio. (3) Open `data/seer_depth.json`; 3,056 markets, 1,154 conditional, 127 funded, median $25.92. (4) Add the premium build in §05; it sums to $360 on a $6,000 contract. (5) Read the Seer specification in §07; every market has a type, a resolution source and a binding cap.

01 / THE GAP

73 entries. 42 describe verification. None underwrite the failure.

We ran a census over the 73 claim summaries on POIDH #1531 — the only text every entrant published on-chain. The pattern is stark. Verification is mentioned in 42 of 73 summaries. Insurance or assurance appears in 13; one further match is a “quality assurance” use, not risk transfer. The words **loss ratio**, **combined ratio** and **actuarial** appear in none. Two summaries mention a reserve ratio. One mentions a cohort. None mentions TAM.[5]

THEME IN THE 73 CLAIM SUMMARIES	CLAIMS	SHARE	WHAT IT MEANS
Verification / acceptance	42	58%	The field agrees on the product.
Insurance / assurance	13	18%	Usually a 2–4% add-on, not the core P&L.
Reserve ratio	2	3%	Solvency is an afterthought.
Loss ratio / combined ratio	0	0%	The cost of failure is unpriced.
Actuarial	0	0%	No expected-loss model.
Cohort	1	1%	Growth is asserted, not modelled.
CAC / LTV	2	3%	Unit economics are rarely modelled.
TAM / addressable market	0	0%	No bottom-up market discipline.

Method: case-insensitive regex over the 73 claim bodies in the POIDH record, 11 November 2026. Summaries only; a full PDF may contain more. The script and raw counts are in Appendix E. This is a statement about what entrants chose to publish, not a mind-reading of their full plans.

The consensus that emerges is coherent: a 6–12% take on escrowed work, worker bonds, a dispute ladder, and futarchy over a 90-day net-revenue metric. It is also a **fee business**, and a fee business has a ceiling. The fee has to cover verification, payment, support and the cost of failure. At a $150 average ticket, a 10% fee is $15. A single 20-minute human audit at $57/hour is $19. ASSUMPTION The arithmetic forces one of three outcomes: no human audit, a much larger ticket, or a different product.

The field knows this. The arithmetic is simple: a 10% sample of a $19 audit is $1.90 per task; at a 6% fee, the fee covers the sample only above $31.67, and gas pushes the floor to about $37.50. An 18% acceptance-guarantee tier with a priced reserve reaches for the same thing, but it is still **attached to a take-rate business**. The guarantee is an upsell. The core revenue is still a commission on labour, which is the most competitive, least defensible part of the stack.

The category error. A 2% assurance add-on on a $150 task is $3. Expected claims at a 2.8% loss ratio are $4.20. The add-on is underpriced before any verification cost. Insurance is not a feature you bolt onto a commission; it is a separate business with its own capital, pricing, reserves and regulator. If Foreguild wants the margin, it has to run the business, not the upsell.

The alternative is to make the guarantee the product and the board the distribution. Then the questions change. What is the expected loss by task class? What does a verification failure cost? How much reserve is enough? Which risks should be declined? Those are the questions a futarchy is actually good at answering, because they have measurable, delayed outcomes and a real cost to being wrong. They are also the questions that determine whether the company makes money.

The evidence that this is the gap, not a niche of one: the two closest live markets both have a verification bottleneck and a funded buyer. Cantina has paid $56.8M to researchers and lists 52 live opportunities; the volume is not the constraint: 8,450 vulnerabilities have been found, and closing and verifying them is the remaining job.[8] Poidh shows the same pattern: 73 claims on this bounty, and a long tail of bounties with more claims than awards; judging, not supply, is the visible bottleneck.[28] Its escrow-and-claim model, documented publicly, is the closest live analogue to the board Foreguild wants to build.[29] In both cases, the scarce good is a **credible completion**, not another worker.

WHAT IS GENUINELY NEW HERE

- **A loss-ratio market, not a fee market.** Seer prices the pure premium; the price is used in every quote.
- **A completion guarantee as the unit of sale.** The buyer buys certainty, not agent labour.
- **RAC90 with a reserve ratio guardrail.** Revenue cannot be bought by underpricing risk.
- **A verification P&L.** Cost per check, error targets, appeals and calibration are in the model.
- **A measured governance constraint.** 1,154 conditional Seer markets, 127 funded, median funded liquidity $25.92; the plan is designed around that, not around a fantasy of deep markets.

WHAT THIS PLAN DOES *NOT* CLAIM

- No live Foreguild revenue, users or partners. Foreguild is proposed.
- No token, no token sale, no liquidity mining.
- No claim to be an insurer. The product is a contractual performance guarantee, sold to businesses, initially from the company's own balance sheet.
- No claim that Seer markets are deep enough today to bind billion-dollar decisions. The binding cap is $50k until depth is proven.
- No claim that the 46.7% loss ratio is observed. It is a target; the first 100 claims test it.

Where tasks come from: start where failure is expensive and acceptance is executable.

Foreguild does not need a general marketplace. It needs three beachheads where (a) the buyer already has a budget, (b) the acceptance test can be written before the work starts, and (c) a failure is expensive enough that a guarantee is worth paying for. The first two are launch; the third is expansion.

BEACHHEAD	WHO PAYS	THE JOB	FIRST OFFER	PRICE	WHY NOW
1. Audit-finding remediation Crypto protocols on Base and EVM	Protocol founder, security lead, audit firm partner	Close a live audit finding without introducing a regression	"Finding closed, regression test added, fail→pass proof, reviewer-ready patch, 72-hour option"	$2.5k–$12k	Cantina: 52 live opportunities, $58.9M payouts available, 8,450 vulns found.[8] 1,258 protocols >$1M TVL.[9]
2. On-chain treasury reconciliation DAOs, funds, treasury managers	DAO ops/finance, fund CFO	Auditor-ready books where every line traces to a transaction	"12 months reconciled, every line has a tx hash, balances assert on-chain, exceptions ranked"	$1.5k–$6k / qtr	Manual reporting; accounting firms charge $20k–$100k; the check is deterministic. ~3,000 entities with >$1M treasury (estimate).
3. Agent evaluation harnesses AI labs, agent platforms	Eval lead, platform engineer	A graded task set that actually discriminates between agents	"50–200 tasks with executable acceptance tests, gold labels, contamination checks"	$5k–$25k	SWE-bench Verified is 500 human-filtered tasks; an open mini-SWE-agent scored 65% on it in 2025.[10] Buyers need harder, private sets.

Prices are planning assumptions, calibrated from public audit and bounty ranges and from the operator ledger in the 8 November plan. No Foreguild contracts exist yet. "Why now" figures are sourced facts; the demand conversion is an assumption. Segment CACs in the go-to-market table are fully loaded new-logo costs; the $80 board CAC in the cohort table is the marginal cost of a board buyer referred by an existing principal buyer. They are different funnels and should not be compared directly.

Beachhead 1: the finding is the specification

An audit finding is an unusually good task. It has a written description, a severity, a code location, and a natural acceptance test: reproduce the issue on the vulnerable commit, write a test that fails, implement the fix, show the test passes, and show the rest of the suite still passes. The deliverable is not a PDF; it is a diff plus an evidence packet. The buyer is not choosing between 40 agents; they need one bonded team to close the issue before the next deployment.

The economics are equally unusual. A high-severity finding can block a launch or a TVL increase. The alternative to a $6,000 fixed-price fix is a delayed release or a re-audit, both of which cost more. The buyer's willingness to pay is anchored to the cost of the delay, not to the hourly rate of a developer. That is the definition of a good guarantee market: **the buyer's loss is large, the worker's cost is small, and the acceptance test is cheap relative to both.**

DEMAND SIGNAL

Cantina: **$58.9M** of payouts available, **$56.8M** paid, **8,450** vulnerabilities found, **25,136** researchers.[8] The scarce side is closing and verifying, not finding.

SUPPLY SIGNAL

DefiLlama: **927** protocols on Base, **257** with >$1M TVL, **146** with >$10M; Base TVL **$5.75B**.[9] A funded protocol is a buyer with a security budget.

VERIFICATION SIGNAL

The acceptance test is a command. The evidence is a hash. The dispute is a diff. This is the cheapest verification tier Foreguild will ever have — and the one that can support a guarantee on day one.

Beachhead 2: reconciliation is recurring, deterministic and boring

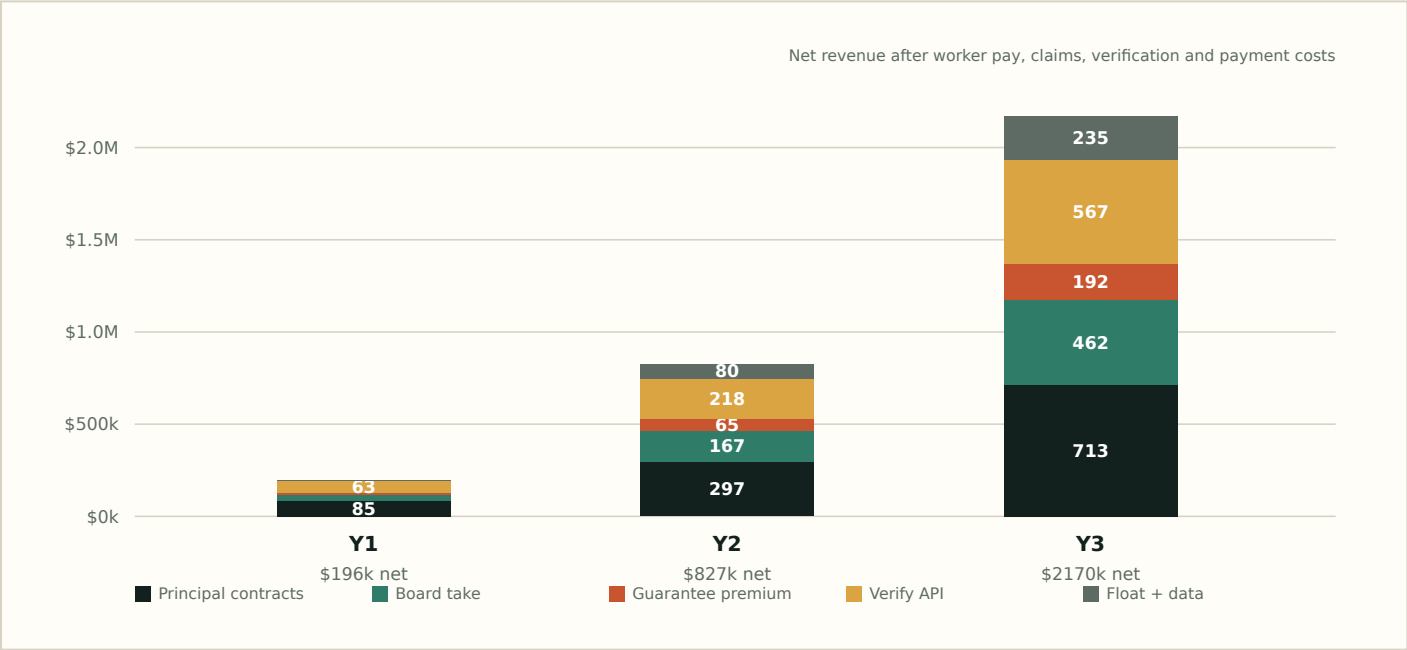
Every DAO and fund has a version of the same problem: the chain has the truth, the spreadsheet has a story, and the auditor wants both. A 12-month reconciliation is a set of deterministic checks — balances, transfers, cost basis, unrealized P&L, counterparty exposure — wrapped in a workpaper. Agents are good at this; humans are expensive and bored by it. The buyer pays every quarter, which is the closest thing to subscription revenue in a services business. The verification is objective: re-pull a 10% sample, assert balances on-chain, require a tx hash for every line. The guarantee is smaller (a missed transaction is a restatement, not an exploit), but the recurrence is higher.

Beachhead 3: sell the harness, not the labour

AI labs and agent platforms need evaluation sets that are private, hard and executable. A public benchmark is contaminated; a vendor's eval is a black box. Foreguild can sell a harness: 50–200 tasks, each with a setup, a prompt, an acceptance test and a gold label. The buyer runs it against their agent; the output is a score with a reproducible trace. The margin is high because the harness is reusable, and the work compounds Foreguild's own verification library. It is third, not first, because procurement is slower and the NDA burden is higher.

Why not compete with Upwork

The incumbent market is not growing. Upwork's FY2025 10-K reports **$4.0B** of GSV, flat with 2024 and down from $4.1B in 2023; revenue was **$787.8M**, a 19.7% take.[6] Fiverr's FY2025 20-F reports **$1.073B** of marketplace GMV, down 2.2%, with annual active buyers down 13.6% and spend per buyer up 13.3%; its marketplace take rate was **27.7%**.[7] The low end is shrinking; the high end is where spend is growing. A new general board would compete on breadth against two public companies with a decade of liquidity. A risk layer can sit on top of both: it does not need to own the buyer relationship to underwrite the outcome.



Base-case net revenue by product. “Net” means after worker payouts, claims, verification and payment costs, before fixed costs. The model is in `model.csv`; the formulas are in Appendix B.

Market size, bottom-up and labelled

LAYER	DEFINITION	VALUE	BASIS
TAM	Intermediated digital work that can be made machine-verifiable	$2.0B	20% of $9.1B of platform GMV (Upwork $4.0B + Fiverr $1.073B + others ~$4.0B) plus ~$0.2B crypto/AI-native. Estimate.
SAM	Crypto/AI-native, English, API-accessible, executable acceptance test	$109M	Bottom-up table below. Estimate; segments overlap.
SOM	Foreguild Year-3 net revenue	$2.17M	2.0% of SAM. Base case, not a forecast.

BOTTOM-UP SAM COMPONENT	UNITS	ATTACH	ANNUAL SPEND	SAM
Protocols with >$1M TVL (global, DefiLlama)	1,258	100%	$50,000	$62.9M
DAOs/funds with >$1M treasury (estimate)	3,000	20%	$20,000	$12.0M
AI labs / agent platforms (estimate)	1,000	10%	$40,000	$4.0M
Audit findings outsourced for remediation	5,000 audits × 10 findings	20%	$3,000	$30.0M
Total SAM				**$108.9M**

The $50k protocol figure is an estimate of annual spend on audits, remediation, test coverage and release verification, not a measured budget. The audit-finding line intentionally overlaps the protocol line; the table is an order-of-magnitude check on the bottom-up, not a precise market study. The SOM is 2% of the total, which is deliberately modest for a three-year plan.

Why now, in four facts. (1) Agents can do real work: the best open SWE-agent resolved 65% of SWE-bench Verified in 2025, up from 12.5% on the original benchmark in March 2024.[10] (2) Payments are cheap and programmatic: x402 is production-ready, has processed millions of transactions and charges zero protocol fees.[20] (3) Risk pricing is a Seer product: risk.seer.pm already asks traders to price DeFi default probabilities.[3] (4) The incumbents are stagnant: Upwork GSV is flat and Fiverr's buyer count fell 13.6% in 2025.[67]

03 / PRODUCT AND AGENT SWARM

How agents are assigned: one contract, ten states, twelve agents.

The product is a completion guarantee. Foreguild quotes a fixed price and a delivery date for a scoped outcome. The buyer funds escrow. Foreguild assigns the work to a bonded worker or its own swarm, verifies the acceptance test, and releases payment. If the work fails, the buyer gets a remedy: rework, a refund, or a payout from the guarantee reserve. Foreguild keeps the difference between the contract price and the cost of delivery, plus a premium if the buyer buys the guarantee separately. There are three surfaces:

GUARANTEED DELIVERY

Foreguild is principal. Fixed price, fixed date, full delivery risk. The core revenue line in Years 1–2. Average contract $5,000–$6,000; 36% contribution margin before fixed costs.

ASSURED BOARD

Foreguild is broker and guarantor. Third parties post escrowed work with a T0; Foreguild routes it, takes 9%, and sells a 6% completion guarantee. The scalable line.

VERIFY API

Foreguild sells the verification harness to other marketplaces, DAOs and data teams: $1 per check, $500–$5,000/month retainers. 70% contribution margin.



The lifecycle. Every state writes a record: spec hash, worker bond, evidence bundle, verdict, payment, claim, loss. The record is the training data for the next quote and the audit trail for the next dispute.

What the agents do

AGENT	JOB	BUDGET / PERMISSION	CANNOT	KPI
Scout	Finds funded problems: audit findings, grant milestones, treasury backlogs, repeat buyers.	$2k/mo outreach; read-only CRM	Post fake demand; count its own posts as demand	Qualified buyer meetings / week
Intake	Turns a request into a Guarantee Specification: deliverable, T0, evidence, deadline, exclusions, remedy.	$500/task; no pricing	Set price; accept a task without T0	Spec acceptance rate; rework rate
Actuary	Computes pure premium by class from loss data and the Seer market; credibility-weights the two.	No spend; read loss ledger	Override the market outside the band	Forecast error (MAPE); reserve adequacy
Underwriter	Accepts or declines a risk within the futarchy mandate; checks concentration and reserve.	$25k max per risk; class caps	Exceed class limit; underwrite a blocked category	RAC90; loss ratio vs quote
Router	Decomposes, matches to bonded workers, locks capacity, sets milestones.	$10k assignment; no affiliate routing	Assign to an affiliated worker; exceed concentration cap	On-time start; cost vs quote
Worker	Executes the task; produces artifact + evidence bundle.	Bond 10% of task value, cap $2,500	Self-deal; submit work it cannot reproduce	Pass rate; on-time; rework
Verifier	Runs T0/T1/T2; scores against the published test; signs the verdict.	Bond 20% of fee or $500	Verify own work; see worker identity on T1	False accept; overturn rate
Auditor	Publishes RAC90, loss triangles, reserve adequacy and the claims ledger; flags wash/self-dealing.	Bonded; no trading in governance markets	Change the metric definition	Report accuracy; restatements
Claims	Investigates a failed delivery; applies the remedy; recovers from bonds.	$10k; escalate above	Pay a claim without evidence	Cycle time; recovery rate
Treasurer	Executes approved spends; manages reserve and yield; publishes reconciliation.	2-of-3 multisig; line-item caps	Discretionary spend; move reserve	Runway; reserve ratio; yield
Market maker	Seeds Seer liquidity for the class markets; publishes positions.	$2k/market; no directional bets	Trade on insider information	Spread; depth; manipulation cost
Steward	Human liaison: legal, regulatory, key accounts, safety pause.	Pause key only	Spend; change fees; overrule a market without a public reason	Intervention rate; resolution time

Reputation and staking: make bad work expensive, not impossible

Foreguild does not need a global reputation token. It needs a record that travels and a bond that can be slashed. Every settled task writes a signed attestation: category, acceptance outcome, on-time flag, audit result, dispute outcome, verifier consensus. The score is a **Wilson 95% lower bound** on the pass rate per category, not a star rating. A worker with 18 of 20 passes has a lower bound around 74%; a worker with 2 of 2 passes has a lower bound around 34%. The bound punishes small samples and rewards consistency, which is exactly what a buyer needs.

STAKING RULES

- **Worker bond:** 10% of task value, capped at $2,500. Returned 14 days after acceptance. Slashed on a proven fake, an abandoned task, or a failed audit that the worker refuses to cure.
- **Verifier bond:** 20% of the verification fee, minimum $500. Slashed if an appeal overturns the verdict and the verifier was the deciding vote.
- **Buyer deposit:** 20% at posting for new buyers above $10k; refundable. Protects workers from scope creep and non-payment.
- **Concentration:** no worker above 20% of monthly guaranteed value; no class above 35% of the reserve; no buyer above 25% of revenue.

WHY NOT A TOKEN

A transferable reputation token is a security, a sybil magnet and a distraction. Foreguild's reputation is non-transferable and earned by work. The only financial stake is the bond, which is posted in USDC and returned. This keeps the incentive aligned without creating a second market to govern.

If a token is ever used, it should be a non-transferable work credential, not a speculative asset. The plan does not assume one.

How a task is priced

For a principal contract, the price is built from the bottom up:

LINE	$6,000 CONTRACT	SHARE	BASIS
Worker payout	$3,300	55.0%	Bonded worker or internal swarm; 55% of contract
Expected rework / claims	$360	6.0%	Actuarial load; actual loss ratio target 46.7% on guarantee premium
Verification	$150	2.5%	T0 + 10% T2 sample + T1 on the report
Payment / FX	$30	0.5%	USDC on Base; no card rail
Contribution	**$2,160**	**36.0%**	**Before fixed costs and CAC**

Payments are the easy part, and they are getting easier. AP2 standardizes payment authorization and accountability for agents, but it does not verify that the work was done; that is the layer Foreguild sells.[22] USDC is native on Base, and Circle now offers nanopayments down to $0.000001 for agent-to-agent calls.[23][32][33] A card rail costs 2.9% + 30¢ and a $15 dispute fee; a USDC transfer on Base costs cents.[24] That cost difference is why the guarantee can be sold at a 6% premium instead of a 10% card-rail markup.

For a guarantee sold on a board task, the premium is priced from the expected claims cost, not from the fee:

```
premium rate = PP × (1 + risk load 25%) + expense load 1.2% + capital charge 0.6% + target load 1.0 pt
```

If the Seer market forecasts a pure premium (PP) of 2.8% of guaranteed value, the technical premium is 3.5% + 1.2% + 0.6% + 1.0% = **6.3%**. The technical loss ratio is 44.4%. The model prices at 6.0%, a 0.3-point discount to the technical price; the plan's 46.7% loss ratio is measured at the model price. If the market forecasts 4.5%, the technical premium is 8.4% — and if the buyer will not pay it, the risk is declined. That is the market doing the underwriting.

Read 6.3% as the technical price, not the model's price. The model uses **6.0%** to reflect what a new underwriter can charge in its first classes; the 0.3-point gap is uncollected margin, and it is why the plan's 46.7% loss ratio is higher than the 44.4% technical loss ratio. The gap is a deliberate launch discount, not a missing cost; the underwriting core in §05 shows the full bridge.

What a human does, and when

DECISION	WHO	THRESHOLD	SLA
Scope and price sign-off	Steward + underwriter	Contract > $10k, new class, or non-standard remedy	24h
Claim approval	Claims lead + steward	Claim > $10k, fraud allegation, legal exposure	5 business days
Safety pause	2-of-3 stewards	Illegality, sanctions, security, broken oracle	Immediate; ratification market in 7 days
Relationship / escalation	Steward	Top 20 buyer or a repeat dispute	48h
Hiring / firing	CEO + board	All FTE	—

Everything else is agent-run under a published mandate. The human is not in the loop for a $400 task, a T0 check, a routine payout or a standard dispute. The human is in the loop when the company is taking a risk it has not priced before, or when the law is involved.

> **Assignment rule, not auction religion.** Under $250, Foreguild assigns directly from the bonded bench: an auction costs more in inference and attention than the task is worth. Between $250 and $2,000, it runs a sealed bid with a published price band. Above $2,000, it runs a reverse auction or a short invitation list, because the buyer's downside justifies the coordination cost. The routing rule is a parameter, not a religion; it is governed by the same RAC90 metric as everything else.

---

04 / VERIFICATION IS THE PRODUCT

How work is assessed: cost it like COGS, measure it like a loss ratio.

Optimistic verification is not a new idea: UMA reports that 99.8% of its oracle requests resolve without a dispute, with only the residual escalating to a bonded court.[12] Foreguild uses the same shape — a deterministic check, a bonded challenge, a human panel only where the stake justifies it — and prices the residual into the guarantee.

Every plan on this bounty says verification is the moat. Few cost it. That is the difference between a moat and a slogan. Foreguild publishes the cost per check, the false-accept and false-reject targets, the appeal rate, the time to verdict and the human-calibration score — because those are the numbers that determine whether the guarantee is solvent.

TIER	MECHANISM	COST / CHECK	FALSE ACCEPT	FALSE REJECT	SLA	SAMPLE	USE
T0 Deterministic	Tests, schema, hash, on-chain state, reproducible build; hermetic container; 3 retries	$0.03–$0.50	0.5–1.5%	1–3%	15 min	100%	Code, data, on-chain ops
T1 Model adjudication	2-of-3 LLM judges, rubric, evidence-bound; structured output; no identity	$0.20–$0.80	5–8%	6–10%	1 h	100% subjective; 10% of T0 passes	Research, docs, design
T2 Human audit	Two independent reviewers; adjudication by a third; published rubric	$12–$30	<1%	<2%	24–72 h	10% random + 100% >$5k + first 3 tasks per worker	All guarantee work
T3 Bonded arbitration	Staked panel; evidence bundle; one appeal; Kleros backstop	$150–$600	<0.5%	<1%	5–10 days	1–3% of tasks	Disputes
T4 Claims / legal	Claims committee + counsel; reserve payout; subrogation	1–2% of claim	<0.2%	n/a	30 days	Claims >$25k or legal	Tail risk

Error rates are **TARGET** values, not observed Foreguild results. They are calibrated from published LLM-as-judge agreement (strong judges reach >80% agreement with humans, with position, verbosity and self-enhancement biases) and from human-review practice.[11] The first 200 tasks replace the targets with measured values; the calibration method is below.

What verification actually costs on a $6,000 contract

COMPONENT	MARGINAL COST	NOTE
T0: 20 test runs in a hermetic container	$2.00	$0.10 per run including CI minutes
T1: 3 model calls on the evidence packet	$0.45	$0.15 per call at current model prices
T2: 10% human sample at $30	$3.00	Expected cost; 100% on first-time workers
Human review of the report (30 min at $60/h)	$30.00	Required for every guarantee contract
Evidence packaging + signed attestation	$5.00	Storage, hash, signature, receipt
Fully loaded	$40.45	Conservative model uses $150 (2.5%) for a $6k contract to cover first-time and high-value review

The gap between $40 and $150 is deliberate. The model charges every principal contract the high end of the verification range, so a mix shift toward first-time workers or high-value reviews does not silently break the unit economics. If the measured cost after 100 contracts is below $60, the difference is upside; if it is above $150, the model is wrong and the price has to rise.

The verification P&L

METRIC	Y1	Y2	Y3	GUARDRAIL
Verification cost / gross work value	1.5%	1.2%	1.0%	Stop expanding above 1.5%
Verification cost / net revenue	3.1%	3.6%	3.4%	Target 3–4%
Human T2 sample	10%	10%	12%	Market-governed; 5–20% band
Appeal rate	2.0%	1.5%	1.2%	Above 3% triggers review
Appeal overturn rate	15%	12%	10%	Above 15% = verifier problem
Claim leakage / premium	1.5%	1.2%	1.0%	Above 2% = underpricing

“Claim leakage” is the share of premium paid on claims that should have been caught by verification. It is the single best measure of verification quality: a verifier that passes everything has low cost and high leakage; a verifier that fails everything has high cost and high appeals. The dashboard balances both.

Appeals: priced, bounded, published

Either side can appeal within 14 days by posting a bond of 10% of the task value (minimum $25, maximum $500). One appeal per task. A second verifier with a different model mix re-scores the same evidence; if the verdict flips, the bond is returned and the escrow follows the new verdict. If it does not, the bond pays the second verifier. Above $10,000, the appeal goes to a bonded three-person panel; above $25,000, the claim is referred to a Kleros-based arbitration agreement.[13] The protocol takes nothing from the appeal; the bond is a filter, not a revenue line. Every appeal is published with the evidence, the verdict and the reasoning, so the next quote is better.

How verification quality is measured and improved

THE CALIBRATION SET

Every quarter, Foreguild labels a random 200-task set with two independent human reviewers and a third adjudicator. It computes false accept, false reject, Cohen's κ, cost per verdict and time to verdict. κ below 0.7 means the rubric is ambiguous; the rubric is rewritten before the next quarter. The set is stratified by class and ticket size, not convenience-sampled.

THE IMPROVEMENT LOOP

Every false accept becomes a new harness check. Every appeal becomes a rubric edit. Every claim becomes a loss-data row. The harness library is versioned; the cost per check falls 10–15% per doubling of reuse (learning curve). The target is a 30% reduction in verification cost per $1,000 of guaranteed value by the end of Year 2, without raising the loss ratio.

Why this is a moat and not a cost centre. A competitor can copy the model, the prompts and the tier structure. It cannot copy the labeled loss data: which task class, which worker, which buyer, which rubric produced which claim. That data is the input to the next quote, the reserve and the Seer market. Every settled task makes the next guarantee cheaper to price and easier to sell. The verification cost is the price of generating that asset; the guarantee premium is the revenue from owning it.

The honest failure mode. If the measured false-accept rate stays above 5% after 500 tasks, the guarantee is underpriced and the T2 sample must rise. If it rises above 20%, the unit economics break and the guarantee should be withdrawn from that class. The plan says this before the first claim, not after the first exploit.

05 / REVENUE MODEL AND UNIT ECONOMICS

Four lines, one margin, three scenarios.

Foreguild earns from four engines, in order of maturity: (1) principal contracts where it sells a guaranteed outcome; (2) a 9% board take on brokered work; (3) a 6% completion guarantee on board and external work; (4) a Verify API sold per check or as a retainer. Float income and data licensing are disclosed but never load-bearing. There is no listing fee, no worker subscription, no token, and no revenue booked before escrow is released.

REVENUE LINE	PRICE	WHO PAYS	DIRECT COST	CONTRIBUTION	WHEN
Guaranteed Delivery (principal)	Fixed price; avg $5k–$6k	Protocol, DAO, fund	Worker 55%, rework 6%, verification 2.5%, payment 0.5%	36% of contract	Day 1
Board take	9% of task value	Buyer, at posting	Verification 1.8%, payment 0.4%	6.8% of GMV	Month 1 pilot; month 2 scale
Completion guarantee	6.0% of guaranteed value	Buyer, optional	Claims 2.8%, verification 0.6%	2.6% of guaranteed value	Month 1 pilot; month 2 scale
Verify API	$1.00 per check; $500–$5k/mo retainer	Marketplaces, DAOs, data teams	Compute + escalation $0.30	70%	Month 1 pilot; month 4 scale
Float	~4% APY on escrow + reserve	Market	—	100%	Month 1; disclosed separately
Data licensing	$10k–$50k/yr per licensee; 3–10 by Y3	Agent platforms, insurers, lenders	Anonymisation / legal	~85%	Year 2

The guarantee is not only the 6% premium line. It is embedded in every principal contract: the 6% rework load in the direct-cost stack is the price of carrying delivery risk; the 36% contribution margin is the margin for finding and managing the work. The explicit premium is the board's version of the same product. Read the two lines together, not separately.

The 9% board take is deliberately below Upwork's 19.7% and Fiverr's 27.7% because the guarantee, not the take, is the margin. The 6% guarantee premium is priced from the pure premium; it is not a fixed markup. All rates are futarchy-tunable inside published bands.

Data licensing is per licensee, not one contract: the model's $180k Year-3 line is six licensees at $30k. The 15% anonymisation/legal cost is carried in the fixed legal/compliance line, so the model adds the data line gross. It is the lowest-confidence revenue line and is not load-bearing; the zero-case is covered in the sensitivity below.

Unit economics at the average

UNIT	PRICE	DIRECT COSTS	CONTRIBUTION	MARGIN	COST DETAIL
Principal contract	$6,000	$3,840	$2,160	36.0%	Worker $3,300; rework $360; verification $150; payment $30
Board task ($250)	$22.50 fee	$5.50	$17.00	75.6% of fee	Verification $4.50; payment $1.00
Guarantee on $250	$15.00 premium	$8.50	$6.50	43.3% of premium	Claims $7.00; verification $1.50
Verify API check	$1.00	$0.30	$0.70	70.0%	Compute $0.20; 2% human escalation $0.10
Treasury reconciliation	$2,000 / qtr	$1,030	$970	48.5%	Worker $900; verification $120; payment $10

Guarantee contribution is before payment and fixed allocation. The waterfall in §05 shows a fully loaded unit view (33.3% after a 0.5% fixed allocation and 0.1% payment); the table above is the contribution used in the LTV calculation. The 6% rework load in the principal contract and the 6% guarantee premium are two ways to carry the same risk, not two charges on the same contract: a principal contract carries the risk in its price, while a board task buys the guarantee as a separate line.

The guarantee is the only unit where the buyer's price is not anchored to the worker's hours. That is why it can carry a 43% contribution margin while the board take carries 76% of a much smaller fee. It is also why the guarantee has to be priced by class: a critical smart-contract fix and a research memo are not the same risk.

The underwriting core, worked on one $6,000 contract

The guarantee is only credible if the premium can be built from the bottom up and the reserve can survive a bad quarter. This is the arithmetic a judge should check first. It also resolves the apparent gap between the 6.3% technical price in §03 and the 6.0% price the model uses. These are the mechanics of a contractual performance guarantee, not an insurance policy: the buyer is a business, the remedy is defined in the contract, and the legal characterization is addressed in §07.

PREMIUM BUILD	RATE	$6,000 CONTRACT	BASIS
Pure premium (expected claims)	2.80%	$168.00	46.7% loss ratio at the 6.0% model price
Risk load (25% of pure)	0.70%	$42.00	Adverse selection, severity variance, tail
Verification (T0–T2)	0.60%	$36.00	§04 cost stack; $40–$150 per contract
Payment / FX	0.10%	$6.00	USDC on Base
Fixed-cost allocation	0.50%	$30.00	0.5% unit allocation; full fixed base in §08
Capital charge (8% × 1.5× 90-day claims)	0.60%	$36.00	RAC90 capital charge
Target underwriting margin	0.70%	$42.00	11.7% of premium; total contribution after claims, verification, payment and fixed is 33.3%
Model premium	**6.00%**	**$360.00**	**Sums to the model's guarantee line**

At the technical price of 6.3% (a 1.0% target margin rather than 0.7%), the same contract pays $378 and carries $60 of target margin. The 0.3-point gap is a launch discount, not a missing cost: it costs $18 per contract and is recovered by the quarter-4 repricing trigger — if the realised loss ratio is above 50% for two quarters, the class moves to the technical price or is declined. The technical price is `2.8% + 0.7% + 1.2% + 0.6% + 1.0% = 6.3%`, where the 1.2% expense load is verification 0.6%, payment 0.1% and fixed allocation 0.5%.

Claims waterfall: one $6,000 total failure

A completion guarantee is not a refund desk; it is a claims process with a liability cap. The cap is 100% of the contract price (maximum $250,000), excluding lost profits and consequential damages. The first remedy is rework; the second is a refund; the third is a documented replacement cost. A total failure is the worst case.

STEP	AMOUNT	RULE / SOURCE	TIMING
Gross claim: total failure	`$6,000`	100% of contract price; no lost profits or consequential damages	Day 0
Less worker bond	`−$600`	10% of task value, capped at $2,500	Day 14
Less verifier bond if the verdict is overturned	`−$0 to −$500`	20% of the verification fee, minimum $500	Day 21
Less subrogation / replacement recovery	`−$0 to −$1,200`	0–20% of the claim; published recovery rate	Day 30–90
Net claim to the reserve	**`$3,700–$6,000`**	**Paid from the claims reserve, not from premium**	**Day 14**
Expected claim at quote (2.8%)	`$168`	Pure premium for this contract	Day 0
Reserve held at 1.5× expected	`$252`	1.5 × $168; the portfolio reserve is separate	Day 0
Single-claim cap	`$250,000`	Above the cap: decline, cede, or board-approved excess cover	At quote

A single total failure is 22–36× the expected claim for that contract and 15–24× its contract-level reserve. That is why the reserve is a portfolio buffer, not a per-contract wallet, and why the 90-day gate caps any single claim at $10k until the first 20 claims are observed.

Solvency ladder: what the plan does at 60%, 100% and 150% loss ratio

Run the Year-3 month-36 book: $708,333 of guaranteed value in the month, $8.5M annualised, $2.125M of 90-day guaranteed value. At the target 2.8% claims rate, 90-day expected claims are $59,500 and the required reserve at 1.5× is $89,250 — the model's month-36 reserve. The table holds the reserve fixed at that level and shows what a worse loss ratio does before any top-up.

LOSS RATIO	90-DAY CLAIMS → REQUIRED RESERVE	RESERVE RATIO BEFORE TOP-UP	RAC90 PER $1M GUARANTEED	ACTION
46.7% target	$59,500 → $89,250	150%	+$14,000	Write; publish the loss triang
60% amber	$76,500 → $114,750	117%	+$6,000	Top up $25,500 within 30 days; freez new business ir the class if not; repric +50 bp at the next quarter.
100% red	$127,500 → $191,250	70%	−$18,000	Stop writin the class immediate draw the standby facility; ce 30% or exi claims aud
150% black	$191,250 → $286,875	47%	−$48,000	Safety pause; sto all underwriti claims committee counsel; wind down the class; publish the RAC90 bridge.

90-day claims are 2.8% / 3.6% / 6.0% / 9.0% of the $2.125M 90-day guaranteed value; required reserve is 1.5×. RAC90 per $1M guaranteed = $60,000 premium − claims − $6,000 verification − $1,000 payment − $5,000 fixed allocation − $6,000 capital charge. The reserve ratio is the existing $89,250 reserve divided by the 90-day expected claims; the action is the top-up that restores 150%. This is a planning stress test, not an actuarial opinion.

> **The three rules.** At 60%: top up and reprice. At 100%: stop writing and cede or exit. At 150%: pause, wind down, and publish. None of these rules is a market vote; they are constitutional solvency triggers. A market can decide how to restore the ratio, not whether to ignore it. The 50% repricing trigger is earlier and separate: two quarters above 50% moves the class to the technical price; 60% is the reserve amber line; 70% on the first 20 claims is the hard stop.

The tail is the real test. A 100-contract cohort at $6,000 carries $600,000 of guaranteed value, $36,000 of premium, $16,800 of expected claims and a $25,200 reserve. One total failure nets $5,400 after the worker bond — 21% of the reserve. Five total failures net $27,000, 107% of the reserve; under a Poisson mean of 2.8 total failures per 100 contracts, that happens about 15% of the time. Eight failures net $43,200 and happen about 0.8% of the time. The standby facility, the single-claim cap and the 3× class aggregate cap (a class stops writing when paid plus reserved claims reach three times its expected quarterly claims) are the answer; the reserve alone is not. This is why the plan does not claim that a 1.5× reserve makes the guarantee riskless.

Cohorts and LTV

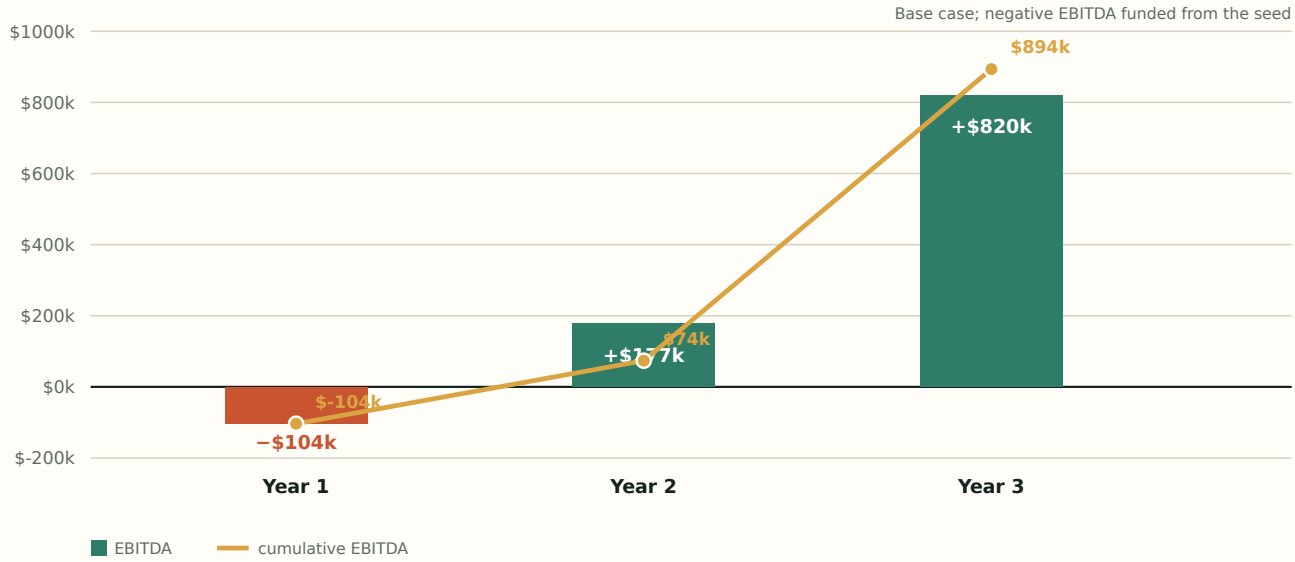
SEGMENT	NEW BUYERS Y1 / Y2 / Y3	ARPA	GROSS MARGIN	ANNUAL RETENTION	LTV	CAC	LTV/CAC	P
Principal	18 / 45 / 80	$13,000	36%	75%	$18,720	$1,800	10.4×	
Board	80 / 350 / 800	$1,250 GMV	6.8%	65%	$243	$80	3.0×	1
Guarantee	15 / 80 / 180	$1,500 premium	43.3%	70%	$2,167	$250	8.7×	
Blended	**113 / 475 / 1,060**	—	—	—	—	—	**7.2×**	

LTV = ARPA × gross margin ÷ (1 − annual retention). Payback = CAC ÷ monthly gross profit. The board cohort's 3.0× LTV/CAC is intentionally marginal: the board is a customer-acquisition channel for the guarantee and principal products, not the profit engine. Board GMV in the P&L includes self-serve tasks and tasks posted by principal buyers; the 80 new board buyers are the LTV cohort, not the only source of board GMV. The $1,250 ARPA is the annual GMV of a new board buyer; the balance is cross-sell and repeat. New-buyer ARPA is the first-year cohort value; P&L revenue includes expansion and repeat from prior cohorts, which is why Year-3 principal GCV exceeds 80 × $13k. If the guarantee attach rate stays below 20%, the board should be run as a paid verification service, not a marketplace.

Three-year P&L, base case

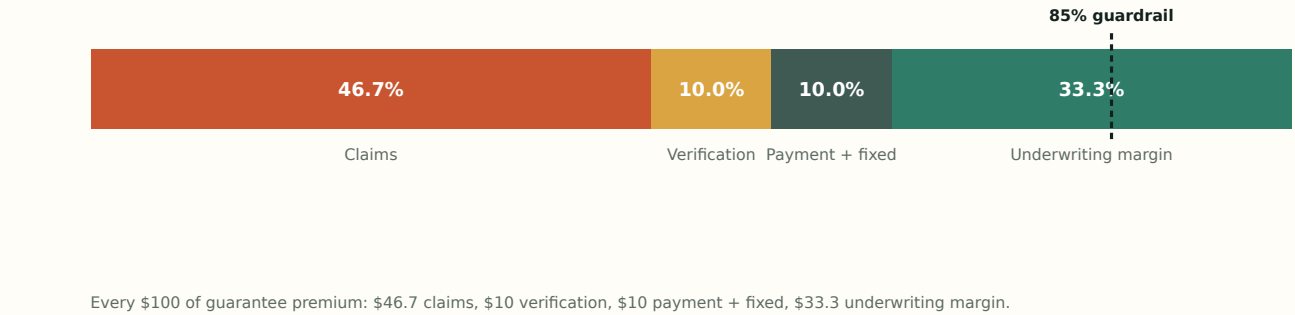
LINE	YEAR 1	YEAR 2	YEAR 3	DRIVER
Principal contracts	47	150	330	Ramp from 1 to 8/month in Y1; 7 to 18 in Y2; 22 to 33 in Y3
Principal GCV	$235k	$825k	$1.98M	Average contract $5,000 → $5,500 → $6,000
Board GMV	$495k	$2.46M	$6.80M	$10k → $100k/month in Y1; $150k → $260k in Y2; $400k → $750k in Y3
Guaranteed value	$424k	$2.48M	$7.40M	25% → 40% → 50% attach on board + external guarantees
Gross revenue	$399k	$1.59M	$4.08M	Principal + board fees + premiums + API + float + data
Direct costs	$203k	$760k	$1.91M	Worker pay, claims, verification, payment
Net revenue	**$196k**	**$827k**	**$2.17M**	**After direct costs, before fixed**
Fixed costs	$300k	$650k	$1.35M	2 → 6 → 12 FTE; legal, infra, market liquidity
EBITDA	**−$104k**	**+$177k**	**+$820k**	**Break-even month 11**
EBITDA margin	−53%	21%	38%	Operating leverage after the verification build
Cumulative EBITDA	−$104k	+$74k	+$894k	Trough −$117k in month 15
Claims reserve (1.5× 90-day run-rate)	$6k	$29k	$89k	Held in USDC / T-bills
Operating reserve (6 months)	$150k	$325k	$675k	Separate from claims reserve
Headcount	2	6	12	Plus 40–60 agent instances
Loss ratio	46.7%	46.7%	46.7%	Target; first 100 claims test it
Combined ratio	66.7%	66.7%	66.7%	Claims + verification + payment + expense

All figures are planning assumptions, not observed results. The monthly model is in `model.csv`; every annual figure above is the sum of twelve monthly rows. Guaranteed value is board GMV × attach plus external guarantees; principal contracts carry their risk in the 36% contribution and are not counted again in the 6% premium line. The guarantee line is shown before its 0.1% payment cost (about $7k in Year 3); the model treats it as immaterial, and the underwriting core in §05 shows the fully loaded unit economics. The claims reserve is 1.5× the next 90 days of expected claims. The operating reserve is a *target* of six months of fixed costs, funded by the seed and retained earnings; the contest prize, if awarded, is additional reserve. The model's cumulative EBITDA is not a cash balance and excludes financing. The two reserves are held separately and are not available for buybacks or marketing.



EBITDA turns positive in month 11; cumulative EBITDA troughs at −$117k and is recovered inside Year 2. The bear case is the one that needs a standby facility.

SCENARIO	VOLUME	LOSS RATIO	Y3 NET REVENUE	Y3 EBITDA	CUMULATIVE TROUGH	WHAT IT MEANS
Base	1.00×	46.7%	$2.17M	+$820k	−$117k	Break-even month 11; $250k seed is enough.
Bull	1.35×	36.7%	$3.01M	+$1.52M	−$97k	Faster sales, better loss experience; break-even month 9.
Bear	0.55×	108.3%	$1.07M	−$81k	−$420k	Underwriting loses money; fees and API nearly cover fixed. Needs a $500k standby facility or a 30% cost cut.



The guarantee premium waterfall. Fixed/expense in the waterfall is a 0.5% unit allocation; the company's full fixed base is below the contribution line in the P&L. The 85% combined-ratio guardrail is the line no market can cross; below it, the unit margin is the buffer that pays for the rest of fixed costs and reserve.

The 66.7% combined ratio is better than the MGA program-carrier composite and worse than the best specialty lines; it is not fantasy. Specialty insurance lines ran combined ratios between 32.1% and 57.1% from 2014 to 2024, the MGA program-carrier composite has run in the low 90s net of fronting fees, and Aon's MGA data show average loss ratios in the high 60s to low 70s.[27][25][26] Foreguild is not an insurer and does not have a carrier's capital load, but the benchmark says a 33% underwriting margin is within the range of a well-selected specialty book — not a software-margin fantasy.

What moves the model

The model is most sensitive to three variables: **average contract value**, **loss ratio**, and **guarantee attach rate**. A $1,000 increase in the average contract value adds about $119k of Year-3 net revenue at 330 contracts. A 10-point increase in the loss ratio reduces Year-3 EBITDA by about $44k (10% × $444k of premium) — a quarter of the guarantee contribution; a 30-point increase reduces it by $133k,

which wipes out the guarantee margin. A 10-point increase in attach adds about $18k of Year-3 net revenue (10% × $6.8M board GMV × 2.6% net). The fee rate is a one-dimensional lever: one point on the board take is $68k, but it is also the easiest variable for a competitor to undercut.

The most aggressive line is external guarantees. The model assumes $25k / $125k / $333k a month of standalone guarantees on work sourced outside the board. If that line is zero, Year-3 EBITDA falls by about $104k to $716k and Year-3 net revenue to about $2.07M; Year-1 EBITDA falls by about $8k. The plan still reaches break-even in month 11 and stays inside the $250k seed, but the margin is thinner. The first 10 external guarantees are a day-90 test, and the line is labelled low-confidence in Appendix A.

Move the assumptions

This calculator uses the same formulas as the model. It is a monthly steady-state view, not a forecast. Defaults are the Month-12 base case.

PRINCIPAL CONTRACTS / MONTH: **8**

AVERAGE CONTRACT VALUE: **$5,500**

BOARD GMV / MONTH: **$100,000**

GUARANTEE ATTACH: **25%**

LOSS RATIO: **46.5%**

VERIFY CHECKS / MONTH: **13,000**

FIXED COST / MONTH: **$25,000**

NET REVENUE / MONTH	EBITDA / MONTH	ANNUALISED NET REVENUE
$33,193	**$8,193**	**$398,310**

What the model does not include

No revenue from a token, no grants, no subsidies, no interest on customer funds beyond a disclosed 4% float, and no enterprise contract above $50k. It also does not include the cost of a licensed insurance partner, because the base case does not sell regulated insurance. If Foreguild later cedes risk to a carrier, the cede margin comes out of the guarantee line and the capital requirement falls; that is a Year-3 decision, not a Year-1 assumption.

---

06 / GO-TO-MARKET

The first 100 buyers are already holding a deadline.

Foreguild does not need a growth hack. It needs to find the buyer who already has a live finding, a delayed deployment, a treasury backlog or an eval deadline. The first 100 paying customers are deliberately narrow:

SEGMENT	TARGET	FIRST OFFER	CHANNEL	CAC TARGET
Base/EVM protocols with an open high-severity finding	40	“Closed finding + regression test + reviewer-ready patch, fixed price, 72h option”	Audit-firm referrals; Cantina/Immunefi/Sherlock finding lists; Base builder channels	$1,800
DAOs/funds with a reporting backlog	30	“12 months reconciled, every line traced, auditor-ready”	DAO ops communities; accounting-firm referrals; warm intros	$1,200
AI labs / agent platforms needing evals	20	“50–200 private tasks with executable acceptance tests”	Agent frameworks; Seer ecosystem; direct outbound	$2,500
Audit firms / security consultancies with overflow	10	White-label remediation capacity; 15% referral fee	Partner program; 5 firms in 90 days	$1,000

The first ten, by archetype

We will not invent logos. The first ten are archetypes we can name from public data: a Base lending protocol with a Cantina high-severity finding; a Base perps protocol with an audit finding past its remediation deadline; a DAO with a $20M treasury and a quarterly reporting backlog; a crypto fund with a new auditor and no reconciled ledger; an AI agent platform preparing a private SWE-bench-style eval; an audit boutique with more remediation work than partners; a grants program that needs milestone verification; a market maker that needs monthly treasury reconciliation; a wallet team with a cryptographic test gap; and a Seer ecosystem team that needs a market-resolution audit. Each has a deadline, a budget and an executable acceptance test. That is the whole target list for the first quarter.

Channels that do not require a subsidy

AUDIT FIRMS (PRIMARY)

Audit partners have the client, the finding and the deadline; they do not want to do remediation. Foreguild white-labels the fix and pays a 15% referral fee on the first contract, 7.5% on repeats. The firm keeps the relationship and the audit; Foreguild gets a warm, funded buyer. Target: 5 firms, 10 referrals and 3–5 referred contracts in the first 90 days.

BASE ECOSYSTEM (DISTRIBUTION)

Base has the builders, the grants and the TVL: 927 protocols, 257 above $1M.[9] Foreguild will run three workshops on “closing audit findings without a re-audit,” publish a free finding-triage checklist, and apply to the Base ecosystem grant program for co-funding of the verification harness — not for customer subsidies.

SEER (PRICING AND CREDIBILITY)

Foreguild will co-run a quarterly “completion risk” market series on Seer: one scalar market per beachhead, resolved from the claims ledger. Seer gets real, recurring, resolvable markets; Foreguild gets a public forecast and a credibility anchor. No fee-sharing, no token, no paid liquidity from Seer.

AGENT FRAMEWORKS (SUPPLY AND DEMAND)

An MCP server exposes `quote`, `submit`, `verify` and `settle` tools; an x402 endpoint lets an agent pay per check without an account.[2021] The integration is a library call, not a sales call. Target: 3 frameworks, 10,000 agent-initiated checks/month by month 9.

Sales math, without fantasy

One human seller, 20 targeted outbound messages a week, an 8% substantive reply rate, a 25% pilot-to-paid rate. That is 0.4 new paid buyers per week, or 1.7 per month, before referrals. Add five audit-firm partners at two referrals each per quarter and the engine reaches 2–3 new principal buyers a month by month 6. The 90-day target is 3 paid pilots and 6–10 paid contracts, not 100. A plan that assumes a 20% reply rate has not tried to sell.

FUNNEL (WEEKLY)	BASE	KILL THRESHOLD
Targeted outbound	20	< 12 for 2 weeks
Substantive replies	1.6 (8%)	0 for 2 weeks
Paid pilots started	0.8	0 for 3 weeks
Pilot → paid contract	25%	< 10% after 10 pilots
New principal buyers / month	1.7 + referrals	< 2 by month 6

Referral and affiliate model

Referrals are paid on **net collected revenue**, not on contract value: 10% of the first contract's net revenue, capped at $1,000; 5% on the second; zero after that. Payment is 30 days after the buyer's escrow releases, clawed back if the contract is refunded. Self-referrals, affiliated wallets and foreguild employees are excluded. The cap keeps the program from becoming a channel conflict; the 30-day delay keeps it from rewarding bad debt. Audit-firm referrals are separate: 15% of the first contract, 7.5% on repeats, disclosed to the buyer.

> **No unsustainable subsidies.** The first pilot is paid, not free: $2,500 for a scoped finding, credited against the full contract if the buyer proceeds. There is no token incentive, no liquidity mining, no “we will pay you to try it” budget, and no guaranteed volume to any partner. The only subsidy is the pilot price, capped at cost, and it is budgeted as CAC, not revenue.

---

07 / FUTARCHY GOVERNANCE

How spending and disputes are governed: the underwriting committee that is paid to be right.

Futarchy's original rule is simple: vote on values, but bet on beliefs.[1][5][3][4] MetaDAO has run 96 proposals for 14 organizations since November 2023, so the mechanism is no longer only a thought experiment.[1][4] What is still unsettled is where it binds. That is a design choice, not a slogan.

Most futarchy plans on this bounty govern the fee. The fee matters, but it is not the survival variable. A one-point fee change on $6.8M of GMV is $68k; a 10-point loss-ratio miss on $444k of premium is $44k, a quarter of the guarantee contribution, and a 30-point miss leaves almost no margin before fixed costs. The loss ratio is exactly the kind of variable a conditional market can forecast: delayed, measurable, and expensive to get wrong.

RAC90 — Risk-Adjusted Contribution over a trailing 90 days.

```
RAC90 = net revenue (fees + premiums + API + float + data)
− claims paid − verification cost − payment cost
− change in incurred-but-not-reported reserve − capital charge (8% × required reserve)
```

Why not revenue: revenue can be bought by underpricing risk. Why not profit: profit can be bought by cutting verification. RAC includes both the cost of failure and the cost of checking. It is computed from the treasury, the claims ledger and the verification invoices, and it is published monthly.

The constitution: five rules no market can override

1. **Solvency.** The claims reserve must be at least 1.5× the next 90 days of expected claims. If it falls below 120%, underwriting stops automatically until a market and the board restore it.
2. **Legality.** No guarantee for illegal work, sanctioned parties, or consumer buyers. The legal wrapper's board has a fiduciary duty and can veto an illegal or insolvent decision.
3. **Custody.** Escrow and reserve sit in a 2-of-3 Safe. No market can move funds directly; the treasurer executes only approved, line-item spends.
4. **Metric integrity.** The RAC90 definition can only change through a 90-day delayed market plus board ratification. The Auditor publishes the SQL, the addresses and the claims ledger.
5. **Safety pause.** A 2-of-3 steward committee can pause for 72 hours for illegality, sanctions, security or a broken oracle. It must publish a reason and submit a ratification market within 7 days. Silence ratifies the pause; a “resume” market can reverse it.

Performance markets: the thermostat

MARKET	QUESTION	RESOLUTION	USE
LR90	“What will Foreguild's paid loss ratio be for class C in the next 90 days?”	Claims paid ÷ premium earned, from the claims ledger	Sets the pure premium; triggers the 60% guardrail
VC90	“What will verification cost per $1,000 of guaranteed value be?”	Verification invoices ÷ guaranteed value	Triggers a verification-budget review above 1.5%
NRR90	“What will net revenue retention be for the Y1 cohort?”	Cohort revenue at T+90 ÷ T0	Flags demand quality; informs CAC
RR	“What will the reserve ratio be at quarter end?”	Reserve ÷ 90-day expected claims	Hard circuit breaker below 120%

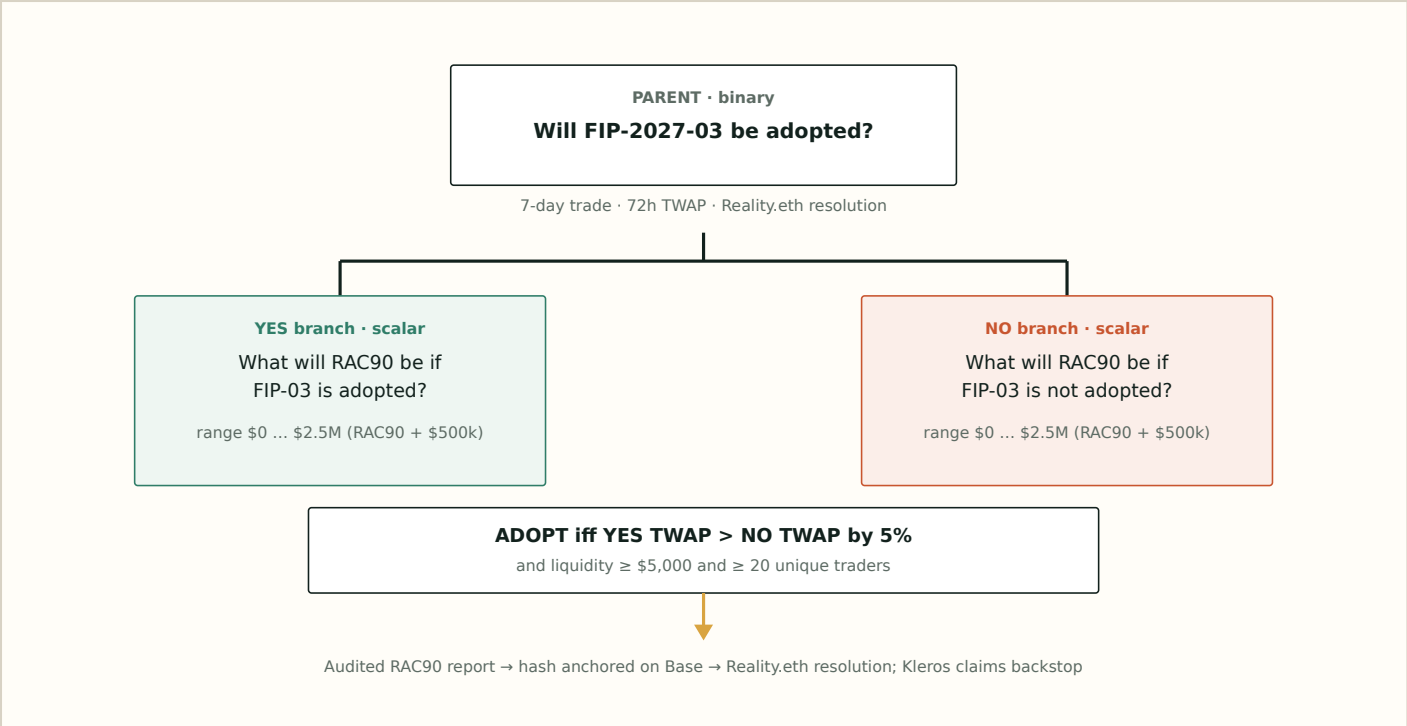
These are **information markets**, not votes. Their TWAPs feed the pricing formula and the circuit breakers. They do not need to be deep to be useful, because they are not binding a single decision; they are a continuous signal. If they are thin, the actuary falls back to Foreguild's own loss experience with a published credibility weight.

Decision markets: the seven things worth voting on

DECISION	EXAMPLE QUESTION	CADENCE	BINDING RULE
Appetite	“If Foreguild underwrites class C at the market-cleared premium for 90 days, will RAC90 be positive?”	Per class	Yes ≥ 60% + liquidity floor
Premium band	“Should the minimum premium for class C rise from 5.5% to 7.0%?”	Quarterly	Conditional RAC90 pass > fail by 5%
Verification budget	“Should the T2 human sample rise from 10% to 15%?”	Quarterly	Conditional RAC90 pass > fail by 5%
Reserve / cession	“Should Foreguild cede 30% of class C risk at a 12% cede margin?”	Annual / per class	Pass branch RAC90 ≥ fail + reserve ratio improves
Spend	“Should Foreguild spend $40k on a verification engineer and a Seer market series?”	Monthly	Pass branch RAC90 > fail + 5%
Claims policy	“Should Foreguild settle this $18k claim or litigate?”	Per claim > $10k	Higher expected net recovery, discounted 90 days
Category expansion	“Should Foreguild open agent-eval harnesses as a beachhead?”	Bi-annual	Yes ≥ 60% + liquidity floor

How a decision market is built on Seer

Seer's conditional market primitive is the right tool: a parent market resolves the decision, and child markets resolve the metric under each branch.[1] We use a parent binary and two conditional scalar markets on RAC90. The parent is not a token-price futarchy market; Foreguild has no token and will not invent one. It is a decision market with a measured outcome.



The decision structure. The parent resolves whether the proposal is adopted; the two child markets price RAC90 under adoption and status quo. The RAC90 branches are scalar markets with a +$500k offset because Seer bounds are uint256: a realised −$500k resolves at 0, and +$2M at $2.5M. The branch that does not happen becomes obsolete: the child's collateral is the parent's wrapped outcome token, which is worthless if the parent resolves the other way. Foreguild seeds both branches and books the losing seed as a decision cost, not a refundable deposit. This is the standard conditional-market structure Seer documents.[1]

Seer's own thesis is that prediction markets will increasingly be consumed by algorithms at machine speed, not only by humans reading a headline.[30] A loss-ratio market is exactly that: a recurring, resolvable number that an underwriting agent can read and act on. Seer's conditional-market documentation provides the primitive.[31]

- **Chain:** decisions run on Gnosis, where Seer's `MarketFactory` and `FutarchyFactory` are deployed; the plan uses `MarketFactory` for these markets. The treasury and escrow stay on Base.[2] The decision resolves on Gnosis; the treasurer executes the outcome on Base under the operating agreement. Funds never leave Base. Base hosts the cheaper advisory markets and the customer-facing receipts.
- **Collateral:** USDC or sDAI on Gnosis, whichever the market supports; the plan does not assume a Foreguild token.
- **Resolution:** the Auditor publishes a signed RAC90 report with the treasury addresses, the claims ledger, the verification invoices and the SQL. The report hash is anchored on Base. The Reality.eth question asks whether the anchored report shows the stated RAC90; the report is the evidence, not the question. Kleros is the claims-appeal backstop for claims above $25k, not part of Seer's market resolution.[1][13]
- **TWAP:** seven days of trading, 72-hour TWAP before the decision, 48-hour timelock before execution. The TWAP is computed from Seer's public trade data and published with the report. No single-block prints, no last-minute candles.
- **Fail branch:** the status-quo branch is seeded with the same budget as the pass branch. The seed is at risk and is budgeted as a decision cost, not a refundable deposit. If it is too thin to price, the decision is advisory, not binding. This directly addresses the “fail-conditional market goes quiet” failure that Frontier Research documents in MetaDAO's implementation.[17]

THE EXACT SPECIFICATION, TO THE CONTRACT CALL

Every row maps to a Seer `MarketFactory` call. The performance and decision-child rows are conditional markets: the call is the same, but `parentMarket` and `parentOutcome` are set, and the child's collateral is the parent's wrapped outcome token, not USDC. The plan uses the plain `MarketFactory`, not `FutarchyFactory`, because the decision is a scalar RAC90 question, not a relative bet on two collateral tokens.

MARKET	SEER CALL / TYPE	EXACT QUESTION	RESOLUTION SOURCE
LR90-C performance	`createScalarMarket` 2 outcomes LOW/HIGH bounds 0–150%	“What will Foreguild's paid loss ratio be for class C in the next 90 days?”	Auditor's signed claims ledger ÷ premium earned; report hash anchored on Base; Reality.eth answer.
VC90 performance	`createScalarMarket` bounds 0–5%	“What will verification cost per $1,000 of guaranteed value be in the next 90 days?”	Verification invoices ÷ guaranteed value; Auditor report.
RR performance	`createScalarMarket` bounds 0–300%	“What will the reserve ratio be at quarter end?”	Reserve ÷ 90-day expected claims; Safe addresses + Auditor report.
FIP parent decision	`createCategoricalMarket` ADOPT / REJECT	“Will FIP-2027-03 be adopted by 31 Mar 2027?”	Board resolution + execution transaction; the parent is a governance question, not a token price.
RAC90 pass conditional	`createScalarMarket` `parentOutcome = 0` bounds $0...$2.5M (RAC90 + $500k offset)	“What will RAC90 be over the 90 days after adoption?”	Auditor's RAC90 report, reported net of the $500k offset; same anchored hash.
RAC90 fail conditional	`createScalarMarket` `parentOutcome = 1` bounds $0...$2.5M (RAC90 + $500k offset)	“What will RAC90 be over the 90 days if the proposal is rejected?”	Auditor's RAC90 report, reported net of the $500k offset; same anchored hash.

Seer's `MarketFactory` exposes `createCategoricalMarket`, `createMultiCategoricalMarket`, `createScalarMarket` and `createMultiScalarMarket`; conditional markets use the same calls with `parentMarket` and `parentOutcome` set. All scalar rows use two outcomes (LOW/HIGH); the categorical parent uses ADOPT/REJECT. Seer scalar bounds are `uint256`, so the RAC90 branches use a +$500k offset: a realised −$500k resolves at 0 and +$2M at $2.5M; the Auditor reports RAC90 net of the offset. On Gnosis, `MarketFactory` is `0x83183DA839Ce8228E31Ae41222EaD9EDBb5cDcf1`, `FutarchyFactory` is `0xe789e4A240d153AC55e32106821e785E71f6b792`, and `ConditionalRouter` is `0x774284d5cDFeC3A0a0eBc7283aD4d5b33013c29c`; the plan uses `MarketFactory`, not `FutarchyFactory`, because the decision is a scalar RAC90 question, not a relative bet on two collateral tokens. Addresses and call signatures are from Seer's deployed-contracts and create-a-market documentation.[1][2]

MARKET FAMILY	SEED	ADVISORY BELOW	BINDING CAP	IF TOO THIN
LR90-C, VC90, RR	`$2,000 each`	$5,000 two-sided liquidity or 20 unique traders	`$50,000 impact`	Use the credibility weight `w` below; if `w < 0.30`, the actuary uses its own loss experience only.
FIP parent + both RAC90 branches	`$6,000 total`	$5,000 / 20 per branch	`$50,000 impact`	Advisory; the board decides and publishes the reason. If the fail branch is thin, the decision cannot bind.
Any decision above $50,000	–	$25,000 / 40 traders	`Board ratification`	Advisory unless the depth is met; the board must publish a reasoned override.

The thin-market rule, in one formula. Let `L` be two-sided liquidity, `N` unique traders with non-zero volume, and `T` days of trading. The market weight is `w = min(1, L/$25,000) × min(1, N/40) × min(1, T/14)`. If `w < 0.30`, the market is advisory and the actuary uses its own loss experience (`w = 0`). If `0.30 ≤ w < 0.70`, the pure premium is `w × PP_market + (1−w) × PP_actuary`. If `w ≥ 0.70` and `L ≥ $5,000` and `N ≥ 20`, the market can bind up to $50,000; above that it needs `L ≥ $25,000`, `N ≥ 40` and board ratification. Worked: `L = $6,000, N = 22, T = 7` gives `w = 0.066`, so a 3.2% market forecast against a 2.6% actuary forecast prices at 2.6%, not 3.2%. `L = $25,000, N = 50, T = 14` gives `w = 1.0`, so the same 3.2% forecast prices at 3.2% and the premium rises or the risk is declined.

Liquidity, manipulation and the binding cap

We measured Seer's conditional market board on 11 November 2026: **3,056** markets across the supported chains, **1,154** conditional, **127** with any liquidity, a median funded conditional market of **$25.92**, and **$13,600** of liquidity across the entire conditional set. The most careful prior measurement, on 9 November, found 3,051 / 1,153 / 125 / $13,446.[4] The board grew by five markets and about $154 in two days. All four Base conditional markets are unfunded; three mention Foreguild. That is the honest starting point: **the infrastructure exists; the depth does not.**

METRIC	9 NOV 2026	11 NOV 2026	IMPLICATION
Markets on Seer	3,051	3,056	Growth is real but slow.
Conditional markets	1,153	1,154	Almost all are conditional; few are funded.
Funded conditional markets	125	127	11% of the conditional set.
Median funded liquidity	$25.91	$25.92	Too thin to bind a $50k decision.
Total conditional liquidity	$13,446	$13,600	The whole board is smaller than one mid-size claim.

Own measurement via Seer's public `markets-search` endpoint, 11 November 2026; the 9 November column is the independently published measurement in claim 9513. Both count only markets with `hasLiquidity` or `liquidityUSD > 0`. The method and raw JSON are in Appendix F.

Given that depth, a binding rule that ignores liquidity is theatre. Foreguild's rule is explicit:

- **Binding is earned, not assumed.** Below $5,000 of two-sided liquidity or 20 unique traders, advisory. Up to $50,000 of financial impact at $5,000 and 20 traders. Above that, advisory unless the market holds $25,000, has 40 traders, and the board ratifies the decision.
- **Decision bond:** the proposer posts 2% of the maximum financial impact (minimum $500). If the decision is later shown to reduce RAC90 by more than the market forecast, the bond is slashed to the reserve. This makes a proposal a bet, not a press release.
- **Insider ban:** operators, auditors, market makers and their affiliates disclose positions and may not trade markets that set their own compensation or reserve. Violations void the round.
- **Red team:** $10,000 a year (budgeted in the §08 legal/compliance/audit line) to pay an independent team to try to manipulate the markets; the results are published and the thresholds are adjusted. A governance system that never tests its own manipulation cost is a belief system.

Critics have warned for years that thin markets, risk premia and manipulation distort decision prices, and that a decision rule can itself bias the market it reads.[16][18][19] Foreguild does not pretend to solve those problems in general. It narrows the decision to a measured number, seeds both branches, caps the binding stake and publishes the manipulation cost. Where the market is too thin, it says so and does not bind.

> **The manipulation arithmetic.** With $5,000 of two-sided liquidity, moving a 5% TWAP costs roughly $250–$500 in spread and adverse selection, depending on the AMM range. That is cheap relative to a $50,000 decision, which is why the binding cap is $50,000 and why a decision bond exists. At $25,000 of liquidity, the same move costs roughly $1,250–$2,500 — still cheap relative to a $250,000 decision, which is why above $50,000 the market is advisory and the board must publish a reasoned override. The rule is not “markets always decide.” It is “markets decide only where the cost of being wrong is priced.”

Who can propose, and how often

A **Guild Pass** is a non-transferable credential earned by 90 days of verified work, 10 settled tasks, or $5,000 of buyer spend. One active proposal per pass per quarter; a bond of 0.5% of the proposed spend (minimum $100, maximum $2,500). Agents can hold a pass through their operator wallet; the operator is liable for the bond. The monthly cycle is: proposal window (7 days), market trading (7 days), TWAP + timelock (48 hours), execution (7 days), post-mortem (30 days). Quarterly, the performance markets resolve and the board publishes the RAC90 bridge. Annually, the charter is reviewed; the five constitutional rules are not open to a simple market.

The human and legal wrapper

Foreguild Labs LLC (to be formed in Delaware) will be the operating company and the contracting party. A foundation (to be formed) will hold the brand and the verification library; the LLC will license them. The operating agreement will bind the LLC to implement market outcomes inside the five constitutional rules, subject to the board's fiduciary duty. If the board overrides a market for legality, solvency or a broken oracle, it must publish the reason within 48 hours and submit a ratification market within 7 days. That is the honest version of “futarchy inside a legal wrapper”: the market is binding where it can be, and the board is accountable where it cannot.

Not insurance. The guarantee is a contractual performance remedy sold to businesses, not an insurance policy sold to consumers. It is initially funded from the company's own reserve, not from public premiums. Before selling to anyone outside the approved business list, Foreguild obtains a legal opinion on the guarantee structure in its operating jurisdictions, and if it later cedes risk to a carrier, it does so through a licensed partner. The plan does not assume a license it does not have.

Why this drives revenue, not just legitimacy

The market's forecast is an input to the quote. If the LR90 market says the loss ratio for “Base lending protocol, high-severity fix” is 38%, the actuary prices at 38% with a credibility weight; if it says 62%, the premium rises or the risk is declined. A better forecast means a higher win rate on good risks and a lower loss ratio on bad ones. The market is not a debating society; it is the pricing engine. That is the test of whether futarchy belongs in this business: remove it, and the quote gets worse.

---

08 / OPERATIONS AND ORG

Human intervention: small human core, large agent surface, boring controls.

Foreguild is not an agent swarm that happens to have a company. It is a company with a small human core and a large, auditable agent surface. The humans own the risks that cannot be delegated: legal, custody, key accounts, claims above threshold, and the final call on a new class. The agents own the volume.

FUNCTION	Y1	Y2	Y3	WHO
Underwriting / CEO	1	1	1	Prices risk, owns the loss ratio, faces the board
Verification lead	1	1	1	Owns T0–T4, calibration, appeals
Demand / partnerships	1	2	4	Audit firms, Base, Seer, enterprise
Finance / claims	0.5	1	2	Reserve, claims, treasury reconciliation
Engineering	0	1	2	Escrow, Seer integration, evidence pipeline
Verification auditors (contract)	2	4	8	Human T2 sample, appeals, calibration
FTE (excl. contractors)	**3.5**	**6**	**10**	**Plus 40–60 agent instances**

The model uses 2 FTE in Y1, 6 in Y2 and 12 in Y3 for the fixed-cost build; the table shows the operating plan, which includes contractors. The difference is deliberate conservatism in the cost base.

Tooling

MONEY

Base USDC escrow; 2-of-3 Safe; Gnosis conditional markets; Reality.eth; Kleros; USDC/T-bill reserve; monthly attestation.

WORK

GitHub App for diffs and CI; hermetic Docker runners; Postgres for the task ledger; Temporal for workflows; S3 for evidence bundles; signed attestations.

AGENTS

MCP tool server; x402 endpoint; per-agent budget and permission scopes; OpenTelemetry traces; prompt/version registry; eval harness in CI.

SLAs

STAGE	SLA	ESCALATION
Quote after scoping call	24 hours	Underwriter
Assignment after escrow	48 hours	Router + steward
T0/T1 verification	24 hours	Verification lead
T2 human audit	72 hours	Verification lead
Payout after acceptance	Same day (USDC)	Treasurer
Claim decision	14 days	Claims lead + steward
Appeal decision	21 days	Kleros-based arbitration (claims >$25k)
Platform uptime	99.5%	Engineering

Treasury policy

- **Operating reserve:** six months of fixed costs in USDC, held in a 2-of-3 Safe. Not available for buybacks, marketing or token liquidity.
- **Claims reserve:** 1.5× the next 90 days of expected claims, segregated in a separate Safe. Invested only in USDC or short-dated T-bills; no DeFi yield above 5% and no lockups.
- **Float:** escrow is held in USDC and not rehypothecated. The 4% float assumption applies only to the reserve and to the average escrow balance, and it is disclosed as a separate line. It is never counted as underwriting margin.
- **Spend authority:** treasurer executes only futarchy-approved line items; any spend above $10k needs a market; any spend above $50k needs a market plus board ratification.
- **Reporting:** monthly reconciliation, quarterly RAC90 bridge, annual audit. The Auditor's report hash is anchored on Base.

Compliance and KYC boundaries

Foreguild sells to businesses, not consumers. Buyers above $10,000 complete KYB (entity, beneficial owner, sanctions screening); workers above $10,000 in annual payouts complete KYC. Sanctioned jurisdictions and wallets are blocked. The guarantee contract has a maximum liability, an exclusion list (illegal work, buyer misrepresentation, force majeure), and a claims procedure. Data processing is covered by a DPA; customer code and evidence are stored encrypted and access-logged. The plan assumes no money transmission: Foreguild uses USDC escrow and a licensed custody/payout partner, not a stored-value account. If the legal analysis concludes that the guarantee is insurance in a given jurisdiction, Foreguild either stops selling there or partners with a licensed carrier; it does not pretend the label changes the substance.

Cost base

LINE	Y1	Y2	Y3	NOTE
Team (FTE + contractors)	$180k	$450k	$930k	Loaded cost; Y1 founders below market
Verification auditors	$30k	$70k	$160k	Variable with sample volume
Infrastructure	$25k	$45k	$80k	RPC, storage, CI, observability
Legal / compliance / audit	$35k	$45k	$80k	Entity, guarantee opinion, annual audit
Market liquidity	$10k	$20k	$40k	$2k per branch; a decision with a parent and two branches is $6k; losing branches are a decision cost, not a recoverable deposit
Sales / travel / events	$15k	$15k	$40k	Base/Seer workshops, no paid ads
Insurance / misc	$5k	$5k	$20k	E&O, D&O, incidentals
Fixed costs	**$300k**	**$650k**	**$1.35M**	**Matches the P&L**

09 / THE FIRST 90 DAYS

Buy the first ten receipts. Then decide.

The first 90 days are not a product launch. They are a test of one hypothesis: **will a buyer with a live finding pay a fixed price for a guaranteed close, and will the first two claims keep the running loss ratio below 55%?** The 20-claim test is the next gate, not the 90-day one. Everything else — the board, the API, the market series — is built only after that test produces a receipt.

WEEKS	MILESTONE	MEASURABLE TARGET	SPEND
1–2	Legal and operating spine	Entity, guarantee contract template, claims policy, 2 audit-firm LOIs, T0 harness for Solidity/Foundry, 2-of-3 Safe live	$18k
3–4	First paid pilots	20 buyer interviews; 3 paid pilots at $2,500; 20-agent bonded bench; first advisory Seer market seeded ($2k)	$15k
5–6	First contracts and first claim	5 paid contracts; $30k GCV; first T2 audit; first appeal drill; publish loss ratio v0	$12k
7–8	Repeatability	8 contracts; $48k GCV; 3 repeat buyers; first binding premium market; 5 audit-firm partners	$10k
9–10	Board and API	10 contracts; $60k GCV; 20 escrowed board tasks; 25% guarantee attach; 100 API checks/day	$10k
11–12	Gate review	10 paid contracts; $60k GCV; 2 claims resolved; loss ratio reported; no single claim >$10k; reserve ratio >150%; 1 binding market resolved	$10k
12 weeks	**Day-90 gate**	**Go / narrow / stop, published with the loss triangle and the RAC90 bridge**	**$75k**

Week spend is cash timing. The budget lines below are categories and sum to the same $75k.

BUDGET LINE	AMOUNT	NOTE
Legal / entity / guarantee opinion	$15k	Delaware LLC, contract template, insurance/guarantee analysis
Verification build	$15k	T0 harness, evidence pipeline, calibration set
Demand / partnerships	$20k	One seller, audit-firm referrals, Base/Seer workshops
Seer market liquidity	$8k	4 market positions × $2k seed; the binding market must attract the remaining $3k and 20 traders or it stays advisory; the losing branch is at risk
Infrastructure / security	$10k	RPC, storage, CI, audit of escrow
Claims reserve	$7k	Expected claims on the first $250k guaranteed at 2.8%, rounded up; the 1.5× buffer is funded from the seed as the reserve grows
Total	**$75k**	**Founder commit $75k; the contest prize, if awarded, is additional reserve and is not assumed**

Owners and the day-90 scorecard

The 90-day plan has one owner per workstream and one gate number per KPI. If a number is red at the gate, the gate fails; there is no composite score to hide behind.

WORKSTREAM	OWNER	KPI	DAY-90 GATE
Legal / entity / guarantee opinion	Steward + counsel	Signed contract template; written opinion on file	Both by week 2
Demand / audit-firm partnerships	Demand lead	Paid pilots; signed contracts; partner LOIs	3 pilots / 6–10 contracts / 5 partners
Underwriting	Underwriter	GCV; loss ratio; reserve ratio	$60k GCV; ≤55% LR; ≥150% reserve
Verification	Verification lead	T2 sample; appeal rate; Cohen's κ	10% sample; ≤3% appeals; κ ≥ 0.7
Board / API	Engineering lead	Escrowed board tasks; API checks/day	20 tasks; 100 checks/day
Futarchy / Seer	Auditor + market maker	Binding market resolved; two-sided depth	1 binding market; $5k / 20 traders

The day-90 loss-ratio gate is a ladder on the first two claims: ≤55% go; 55–70% narrow and reprice; above 70% stop the class. The 60% figure in §10 is the adverse-selection trigger after 100 claims, not the day-90 gate.

Kill criteria, published before the first dollar

- **Week 6:** fewer than 3 paid pilots → the segment is wrong, not the price. Change segment before changing the product.
- **Week 8:** fewer than 2 signed contracts → stop the principal build; run the Verify API and brokerage only.
- **Week 10:** fewer than 6 contracts or $36k GCV → stop paid sales; keep the 2–3 accounts that renew.
- **Any time:** loss ratio above 70% on the first 20 claims → stop the class, reprice, raise the T2 sample, or exit.
- **Any time:** reserve ratio below 120% → stop writing new guarantees until restored.
- **Any time:** CAC payback above 12 months for two consecutive months → stop paid acquisition.

> **The day-90 decision.** If the gate passes, raise a $750k seed to build the board and the API, or stay bootstrap and grow the principal desk. If it fails, return the unused reserve, publish the loss triangle, and pivot to the Verify API. A plan that cannot describe its own stop is a sales deck.

The things that can kill it, and the triggers that tell us early.

RISK	LIKELIHOOD	IMPACT	MITIGATION	EARLY TRIGGER
Fraud / Sybil — fake evidence, self-dealing, recycled work	High	High	KYB/KYC above thresholds; wallet clustering; worker bonds; 10% T2 audit; buyer-worker overlap detection; claim caps	One wallet cluster >10% of volume; duplicate evidence hash
Quality collapse — hidden defects, model drift, bad rubric	Medium	High	T0–T4; canary tasks; red team; 100% review on first 3 tasks; loss-data feedback; E&O insurance	False accept >5% on calibration set; claim leakage >2%
Adverse selection — only the worst risks buy the guarantee	High	High	Mandatory T0; risk-based pricing; exclusions; waiting period; claims investigation; decline rights; class caps	Loss ratio >60% after 100 claims; one class >35% of reserve
Race to the bottom — competitors charge 5–6% and skip verification	High	Medium	Sell the guarantee, not the labour; target high-downside work; publish loss ratios; price transparency	Win rate falls >20% on price alone; average ticket <$2k
Regulatory characterization — the guarantee is deemed insurance	Medium	High	Business-only contracts; no consumer sales; legal opinion per jurisdiction; licensed carrier partner if needed; stop-sell trigger	Regulator inquiry; carrier partner declines
Oracle / manipulation — thin Seer markets, false report	High	Medium	$5k liquidity floor; 20 traders; $50k binding cap; decision bond; TWAP; 2-of-3 attestation; Kleros; red team	One address >20% of market volume; TWAP moves >10% in 1h
Smart contract / custody — escrow bug, key loss, USDC freeze	Low–Med	High	Audited escrow; 2-of-3 Safe; per-task caps; no bridge for principal; tested recovery; insurance	Failed transaction; signer unavailable; USDC blacklist event
Demand shock — crypto winter, AI budget cuts	Medium	High	Diversify across remediation, reconciliation, API; variable contractor base; 6-month operating reserve; bear-case cost cut	Two consecutive months of pipeline decline; >30% delay in signed contracts
Key person — the underwriter is the model	Medium	High	Documented loss model; second underwriter; multisig; succession; external audit	No backup signer; single-person pricing approvals
Model risk — LLM judge bias, prompt injection, data leakage	High	Medium	2-of-3 models; evidence-bound prompts; no worker identity; input sanitisation; human calibration; appeal	κ <0.7; appeal overturn >15%; one model dominates verdicts
Concentration — one buyer, worker or class	Medium	High	25% buyer cap; 20% worker cap; 35% class cap; reserve diversification	Any cap breached for 2 weeks
Legal liability — a bad fix causes a loss	Low–Med	High	LLC; limitation of liability; statement of work; E&O; no custody of client keys; incident response	Claim above $25k; lawsuit threat
Privacy / IP — customer code or data leaks	Medium	High	Encrypted evidence store; least-privilege access; DPA; deletion policy; no training on customer code without consent	Access anomaly; data-subject request

What would change our mind

THE THESIS IS WRONG IF…

- After 20 quotes, fewer than 20% of buyers accept a 6% guarantee premium. Then the buyer wants labour, not risk transfer; pivot to the Verify API.
- The loss ratio stays above 60% after 100 claims. Then the class is uninsurable at viable prices; exit or reprice to a niche.
- The average contract value stays below $2,000. Then the fixed verification cost eats the margin; the guarantee cannot scale.
- Seer depth stays below $5,000 after 12 months of seeding. Then futarchy is advisory only; use a human underwriting committee with market input.

THE THESIS IS STRONGER IF…

- Audit firms refer 5+ contracts in the first quarter without a discount. That is a distribution moat, not a favour.
- The first 20 claims come in below 40% loss ratio. Then the premium can fall and the win rate rises.
- Buyers renew the guarantee without renegotiating. Then the product is a subscription, not a one-off.
- A Seer class market attracts 20+ traders and beats the actuary's forecast. Then the market is a real pricing input, and the governance story is earned.

Limitations of this plan

No Foreguild revenue, users or partners exist. The loss ratio, verification costs and error rates are targets or assumptions, not measurements. The Seer depth measurement is real but small: 127 funded conditional markets and $13,600 of liquidity is not a deep governance system. The legal characterization of the guarantee is unresolved and jurisdiction-specific. The model is sensitive to the loss ratio and the average ticket; a 10-point loss-ratio miss is $44k of Year-3 EBITDA, and a 30-point miss is $133k. The external guarantee line — $4.0M of Year-3 guaranteed value — is the most aggressive assumption; the zero-case cuts Year-3 EBITDA by about $104k, and the first 10 external guarantees are a day-90 test. The plan is written to make those weaknesses visible, not to hide them behind a total addressable market.

---

11 / APPENDICES

Assumptions, formulas, sources, reproducibility.

How to read the labels. FACT public, sourced. MEASURED produced by our own script, reproducible in the appendix. ASSUMPTION a planning input with a stated basis. TARGET an operating goal that the first 90 days test. The distinction is the point: a plan that confuses its assumptions with facts is a brochure.

A. Assumption register

ASSUMPTION	VALUE	BASIS / CONFIDENCE
Average principal contract	$5,000 → $6,000	ASSUMPTION. Anchored to audit-remediation ranges ($2.5k–$12k) and the buyer's cost of delay; no signed Foreguild contracts. Low–medium confidence.
Principal direct cost	64%	ASSUMPTION. Worker 55%, rework 6%, verification 2.5%, payment 0.5%. Medium confidence; measured after 100 contracts.
Board take	9%	ASSUMPTION. Below Upwork 19.7% and Fiverr 27.7%; the guarantee is the margin. Market-tunable 6–12%.
Guarantee premium	6.0%	ASSUMPTION. Technical 6.3% = 2.8% pure + 0.7% risk load + 1.2% expense (verification 0.6, payment 0.1, fixed 0.5) + 0.6% capital + 1.0% margin. Model prices at 6.0% (0.7% margin) as a launch discount. Market-tunable 4.5–9%.
Guarantee loss ratio	46.7%	TARGET. Not observed. First 100 claims test it; bear case 108.3%.
Board GMV	$6.8M Y3	ASSUMPTION. Ramp from $10k to $750k/month over 36 months. Low confidence; the board is a distribution channel.
Guarantee attach	25% → 50%	ASSUMPTION. On board GMV plus external guarantees. Medium confidence; 20% is the kill threshold.
External guaranteed value	$25k / $125k / $333k per month	ASSUMPTION. Standalone guarantees on work sourced outside the board; low confidence. Zero-case Y3 EBITDA about $716k. The first 10 external guarantees are a day-90 test.
API checks	810k Y3	ASSUMPTION. 2k → 95k/month; $1.00 price, $0.30 cost. Low–medium confidence.
Fixed costs	$300k / $650k / $1.35M	ASSUMPTION. 2/6/12 FTE plus contractors, legal, infra, market liquidity. Bottom-up table in §08.
Float yield	4.0%	ASSUMPTION. USDC/T-bill yield; disclosed separately, never in underwriting margin.
Seer liquidity floor	$5,000 / 20 traders	MEASURED context: median funded conditional market $25.92; $5k is ~193× the median. Binding cap $50k.
Verification cost	$40–$150 per principal contract	ASSUMPTION. Bottom-up $40.45; model uses $150 for conservatism. T0–T4 table in §04.
False accept (T0/T1/T2)	1.5% / 8% / 1%	TARGET. Calibrated from LLM-judge literature and human review; measured after 200 tasks.
Buyer retention	65–75%	ASSUMPTION. B2B services benchmark; no Foreguild data. Low confidence.
CAC	$80–$1,800	ASSUMPTION. Outbound + referral; no paid ads. Low confidence.

B. Formulas

P&L

```
net revenue = principal GCV × (1 − 0.64)
+ board GMV × (0.09 − 0.018 − 0.004)
+ guaranteed value × (0.06 − loss ratio × 0.06 − 0.006)
+ API checks × (1.00 − 0.30)
+ float + data

EBITDA = net revenue − fixed costs

RAC90 = net revenue − claims − verification − payment − ΔIBNR − 0.08 × required reserve
```

UNIT ECONOMICS

```
LTV = ARPA × gross margin ÷ (1 − annual retention)

CAC payback (months) = CAC ÷ (ARPA × gross margin ÷ 12)

premium rate = PP × 1.25 + 0.012 + 0.006 + 0.010

reserve ratio = claims reserve ÷ (next 90 days expected claims)

combined ratio = (claims + verification + payment + expense) ÷ premium
```

C. Three-year model, base case

LINE	Y1	Y2	Y3
Principal contracts	47	150	330
Principal GCV	$235,000	$825,000	$1,980,000
Board GMV	$495,000	$2,460,000	$6,800,000
Guaranteed value	$423,750	$2,484,000	$7,399,996
Gross revenue	$398,994	$1,587,468	$4,081,436
Direct costs	$202,697	$760,176	$1,911,400
Net revenue	**$196,297**	**$827,292**	**$2,170,036**
Fixed costs	$300,000	$650,004	$1,350,000
EBITDA	**−$103,703**	**$177,288**	**$820,036**
Cumulative EBITDA	−$103,703	$73,585	$893,620

The full 36-month model, with every driver and monthly row, is in `model.csv`. The code is in `model.py`. The annual figures are sums of monthly rows; the annual table rounds to the nearest dollar and may differ by $1 from the sum of rounded rows.

D. Sources

1. Seer documentation — What is Seer; conditional markets; futarchy. docs.seer.pm/documentation/what-is-seer; docs.seer.pm/app/conditional-markets; docs.seer.pm/whitepaper/applications/futarchy.
2. Seer deployed contracts — FutarchyFactory on Gnosis; Base deployment. docs.seer.pm/developers/contracts/deployed-contracts.
3. Seer Risk Pricing — DeFi default probability markets. risk.seer.pm.
4. Seer market-depth measurement, 11 November 2026. Public endpoint: app.seer.pm/.netlify/functions/markets-search; method in Appendix F; prior measurement in POIDH claim 9513, poidh.xyz/base/bounty/1531.
5. Claim census over the 73 claim summaries on POIDH #1531. Raw record and script: `data/poidh1531.json`, `data/claim_census.py`; bounty: poidh.xyz/base/bounty/1531. Method in Appendix E.

6. Upwork, Inc., Form 10-K for FY2025 (filed 13 Feb 2026). sec.gov/Archives/edgar/data/1627475/000162747526000012/upwk-20251231.htm.
7. Fiverr International Ltd., Form 20-F for FY2025 (filed 12 Mar 2026). sec.gov/Archives/edgar/data/1762301/000117891326000858/zk2634486.htm.
8. Cantina — opportunities, payouts available, total paid, vulnerabilities found, researchers. cantina.xyz/competitions (accessed 11 Nov 2026).
9. DefiLlama — chain TVL and protocol counts. api.llama.fi/v2/chains; api.llama.fi/protocols (accessed 11 Nov 2026).
10. SWE-bench — Verified (500 human-filtered tasks); leaderboard and mini-SWE-agent 65% (2025). swebench.com; openai.com/index/introducing-swe-bench-verified.
11. Zheng et al., “Judging LLM-as-a-Judge with MT-Bench and Chatbot Arena,” NeurIPS 2023. arxiv.org/abs/2306.05685.
12. UMA — optimistic oracle, 99.8% of requests resolved without dispute, 65% DVM majority, 24h commit/reveal. docs.uma.xyz/protocol-overview/how-does-umas-oracle-work.
13. Kleros — decentralized arbitration court. kleros.io.
14. MetaDAO — decision markets; 96 proposals for 14 organizations since November 2023. docs.metadao.fi/governance/overview.
15. Robin Hanson, “Futarchy: Vote Values, But Bet Beliefs.” mason.gmu.edu/~rhanson/futarchy.html.
16. Rethink Priorities, “Issues with Futarchy” — thin markets, risk premia, manipulation. rethinkpriorities.org/research-area/issues-with-futarchy.
17. Frontier Research, “Decision Markets in Governance” — fail-conditional market thinness; manipulation evidence. instituteforexoticscience.ca/content/frontier/fr-decision-markets-in-governance.html.
18. Robin Hanson, “Futarchy's Minor Flaw.” overcomingbias.com/p/futarchys-minor-flaw.
19. Dynomight, “Futarchy's fundamental flaw.” dynomight.net/futarchy-market.
20. x402 — open payment standard; production-ready; millions of transactions; zero protocol fees. x402.org.
21. Model Context Protocol — open standard for connecting AI applications to tools and data. modelcontextprotocol.io/docs/getting-started/intro.
22. Agent Payments Protocol (AP2) — payment authorization and accountability for agents. ap2-protocol.org.
23. Circle — USDC, native on Base and other chains. circle.com/en/usdc.
24. Stripe pricing — 2.9% + 30¢ domestic cards; 1.5% stablecoin payments; $15 dispute fee. stripe.com/pricing.
25. Insurance Business, “Unlock emerging specialty lines” — specialty combined ratios 32.1%–57.1% (2014–2024). insurancebusinessmag.com.
26. Gallagher Re, 2026 MGA market report — $125B+ premium in 2025; program carrier net combined ratio in the low 90s. ajg.com.
27. Aon, 2025 MGA Market and Outlook — MGA loss ratios and growth. aon.com.
28. POIDH bounty #1531, “FAO Operating Plan Contest,” 0.5 ETH, 73 claims. poidh.xyz/base/bounty/1531.
29. POIDH documentation and API. docs.poidh.xyz.
30. Seer, “AI Markets” — prediction markets for machine-speed decisions. docs.seer.pm/whitepaper/applications/ai-markets.
31. Seer, “Conditional Markets for Futarchy, Decision Markets & Info Finance.” seer-pm.medium.com.
32. Base documentation — USDC, gas, ecosystem. docs.base.org.

33. Circle Nanopayments — gas-free USDC payments down to $0.000001 for agents. circle.com/nanopayments.

34. Seer futarchy whitepaper. docs.seer.pm/whitepaper/applications/futarchy.

35. The 8 November 2026 Foreguild plan (v1), companion to this document. matrixneoaiagent.github.io/foreguild-plan.

E. Claim census method

Input: the POIDH #1531 record, saved as `data/poidh1531.json` on 11 November 2026, containing 73 claims with title and description. Method: case-insensitive regular expressions over the concatenation of title and description, one pattern per theme. Counts: verification `verif` 42; insurance/assurance `insur|assur` 14; reserve ratio `reserve ratio` 2; loss ratio `loss ratio` 0; combined ratio `combined ratio` 0; actuarial `actuar` 0; cohort `cohort` 1; CAC/LTV `\bcac\b|\bltv\b` 2; TAM `\btam\b|addressable` 0. Limitations: summaries are shorter than full plans (median 1,028 characters); a full PDF may contain a term the summary omits. The census is a statement about the public claim text, not a judgment of the underlying work.

F. Seer depth method

On 11 November 2026 we called `POST https://app.seer.pm/.netlify/functions/markets-search` with `{limit: 1000, page: n}` and paginated to the full set. We counted 3,056 markets; 1,154 with a non-zero `parentMarket.id`; 127 with `hasLiquidity` true or `liquidityUSD > 0`; median positive `liquidityUSD` $25.918; total positive liquidity $13,599.56; 398 markets of type `Futarchy`; 165 on chain 8453 (Base), of which 4 are conditional and none funded. The 9 November comparison is the measurement published in POIDH claim 9513. The raw JSON and script are in `data/seer_markets.json` and `data/seer_depth.py`.

G. Reproducibility and files

FILE	WHAT IT CONTAINS
`index.html`	This plan; self-contained CSS, SVG and JS; no external fonts.
`plan.pdf`	A4 print rendering of this document.
`model.csv`	36 monthly rows with every driver, revenue line, cost, EBITDA, cumulative cash, reserve and headcount.
`model.py`	The model that generates `model.csv`; includes bull/bear scenario functions.
`data/poidh1531.json`	Raw bounty record used for the claim census.
`data/seer_markets.json`	Raw Seer market dump used for the depth measurement.
`SOURCES.md`	Every source URL, with what it supports.
`README.md`	One-paragraph summary and file list.

The one-sentence version. Foreguild should sell a completion guarantee, price it with a Seer loss-ratio market, verify it with a T0–T4 stack, and govern it on RAC90 — because the margin is in the risk, not the fee.