

Verification Without Known Answers: The Economics of Reusable Judgment

Christian Roessler

Department of Economics; California State University, East Bay

Abstract

In information production, verification solves one hidden-action problem but creates another: reviewers must themselves be induced to exert effort, yet their expertise is the scarce input that verification requires. I show how concealed synthetic errors can break this recursion by rewarding careful review even when correct answers are unknown. Resolved cases can then improve both future assessment and production, allowing costly judgment to be reused. Under sufficient reuse, consistent learning, and vanishing recurrent overhead, quality can approach perfection while average resource cost approaches that of a single routine execution.

Keywords: Anti-Gold; recursive agency; eligibility incentives; data curation; learning; endogenous production layers.

JEL codes: D23; D82; D83; D86; L23; O33.

1 Introduction

The cost of reliable information depends not only on the judgment required to produce it, but also on how widely that judgment can be reused. Obtaining the judgment is itself an agency problem when reviewers' effort is hidden and correct answers are unknown. I study how verification incentives and the reuse of judgments jointly determine the cost and organization of information production.

Resolving a case can create two kinds of savings. Better assessment can sustain production incentives with less costly review. Better execution can reduce the resources needed to meet a given quality target. Verification can therefore be an investment in both supervision and production, rather than merely an expense incurred to inspect current output. Its return depends on how many later decisions can benefit from the knowledge it creates. These gains depend on usable resolutions of ordinary cases.

Verification incentives exploit the difference between knowing that an answer is wrong and knowing what the answer should be. Concealed synthetic errors, called *Anti-Gold*, can make informative review rewardable without a correct reference answer. The checks must reward selective investigation that is also useful on ordinary work.

Production and verification remain economically linked even when different workers perform them. Reviewers can be rewarded directly or through access to future paid work. In

the latter case, the incentive comes from the expected net return to assignments that will actually be offered. Production rewards then support not only production effort but also the review on which its assessment depends. Cutting those rewards can undermine verification even when production would remain incentive-compatible under the intended review policy. Compensation and assignment must therefore be chosen jointly across the two roles.

Learning also changes which resource layers are worth maintaining. A task may cease to require the most expensive resources before routine resources can handle it. Intermediate layers capture savings during this interval, provided that their use across present and future work justifies their setup and maintenance costs. Holding learning speed fixed, more usable resolutions can move additional work into this intermediate range and support more layers. Faster progress toward routine execution can instead shorten the interval and reduce the demand for layers. Optimal depth thus depends on the path of learning, not only on its eventual outcome.

The limiting possibility is Gold-like quality at Bronze-like cost. Gold denotes fidelity to a source and its representation requirements, not the truth of the source itself. Bronze denotes the resource cost of one routine execution. If learning makes increasingly strict quality targets feasible, reuse amortizes preparation, and recurrent overhead per output vanishes, average expected fidelity loss can approach zero while total resource cost per output approaches Bronze. Production and verification must remain incentive-compatible throughout. The cost condition applies to the total resources used by the coordinating organization and workers, including repeated production and continuing review.

Section 2 relates the argument to the literature. Section 3 defines the model, and Sections 4–6 develop verification and production incentives. Sections 7–9 analyze learning, resource layers, and the quality–cost frontier. Sections 10 and 11 consider architecture-dependent learning, identification, and scientific extraction.

2 Connections to the literature

The paper links the incentives needed to obtain informative judgments to the savings available when those judgments can be reused. This connects incentive contracting and information elicitation to learning and knowledge hierarchies. Work on synthetic quality controls supplies a way to induce review without a complete correct reference. The economic question is how inducing that review changes the future cost and organization of production.

Incentives for informative judgment. Incentive contracting studies how observable performance and future rents sustain effort. Holmström (1979) establishes the value of signals informative about hidden action, while Holmström and Milgrom (1991) explain how unequal measurability can direct effort away from poorly measured tasks. Both insights matter here: a review is an incentive signal that requires effort to produce, while a synthetic score may leave important parts of that effort unmeasured. Eligibility rewards also use the disciplinary value of future rents studied by Shapiro and Stiglitz (1984). Here those rents arise from subsequent production assignments, so reducing production rewards can undermine the review needed to assess production.

Information elicitation asks how to reward informed reports when correctness cannot be checked directly. Miller et al. (2005) use correlations between reports to support truthful reporting and effort. Dasgupta and Ghosh (2013) use multiple tasks to penalize low-effort agreement when proficiency depends on effort. Gao et al. (2019) show that, under their conditions, simple checks against limited trusted evaluations can provide stronger incentive guarantees than more elaborate peer mechanisms. These results make the supply of trustworthy evaluations economically consequential. The present model endogenizes that supply by studying incentives for the workers who produce the assessments on which production rewards depend.

Synthetic quality controls. Synthetic quality control uses test cases with known properties to assess work whose ordinary quality is difficult to observe. The software-testing literature explains why detecting a violation need not require a complete correct answer (Barr et al., 2015). Oleson et al. (2011) generate quality-control tasks from modified data, reducing reliance on manually prepared references. Checco et al. (2020) study how workers can identify quality-control questions. In an AI-assisted review experiment, Fang et al. (2026) find that low-frequency injection reduces natural-error detection whereas higher-frequency injection improves it. The present model uses this approach to discipline strategic review. It permits justified negative labels without a certified starting answer and requires the rewarded behavior to be useful on ordinary work. The contribution is an account of when synthetic checks support the judgment needed for production and learning.

Learning and knowledge hierarchies. The literature on knowledge hierarchies explains how organizations economize on expertise. Garicano (2000) assigns common problems to production workers and reserves specialists for difficult cases. Caliendo and Rossi-Hansberg (2012) link the number of layers to the scale of demand. These insights motivate routing each task to the least costly resources that can meet its quality target. Here the workload supporting a layer includes future uses of learned judgments. An intermediate layer can be worthwhile because tasks spend time between expert-intensive and routine production.

Learning by doing makes experience a source of productivity (Arrow, 1962), while Garicano and Rossi-Hansberg (2012) link knowledge accumulation to the development of organizations that exploit technologies. Here usable case resolutions are the learning input, and their production remains subject to incentive constraints. Reuse can improve assessment as well as execution. Holding learning speed fixed, more resolutions can support more intermediate layers; faster routinization can instead make them less useful. This distinguishes the supply of reusable knowledge from the speed at which tasks become routine.

The paper contributes to knowledge-based theories of organization by making the production of reusable judgment an incentive problem. A judgment can lower future execution costs and the cost of sustaining production incentives, so its value depends on the work over which both gains can be realized. Under consistent learning, sufficient reuse, and vanishing recurrent overhead, these gains can bring average expected fidelity loss toward zero at a total resource cost per output approaching routine production cost.

3 Primitives, implementation, and resource accounting

A curator commissions records that represent source information under a fixed protocol. Workers produce the records and review them; their inputs and the correctness of ordinary output are not directly observable. The curator chooses resources, assignments, and compensation to minimize total resource cost subject to a fidelity target. The model specifies quality and feasible resources, workers' choices and assessments, the reuse of resolved cases, and the constraints on the curator's plan. Table 1 collects the main notation.

3.1 Outputs and source fidelity

At date t , a task has observable characteristics $r \sim P_t$ and latent state θ . The characteristics record what is known about its difficulty, risk, omission exposure, and potential for reuse. The representation protocol specifies a finite universe of candidate items $\mathcal{C}_r$, including items that a producer may fail to discover. A released record is $Y \subseteq \mathcal{C}_r$; the nonempty set $\mathcal{G}_r(\theta) \subseteq 2^{\mathcal{C}_r}$ contains all records that faithfully represent the source under the protocol. Several encodings may be admissible. Gold is this quality objective, not an observed reference or a production method.

For an admissible record G , let $w_{rc}^F, w_{rc}^O \geq 0$ weight unsupported inclusions and required omissions. Define

$$F_w(Y, G) = \sum_{c \in Y \setminus G} w_{rc}^F, \quad O_w(Y, G) = \sum_{c \in G \setminus Y} w_{rc}^O.$$

Choose loss weights $\lambda = (\lambda_F, \lambda_O) \geq 0$ before observing output. Source-fidelity loss is

$$d_\lambda(Y, \mathcal{G}_r(\theta)) = \min_{G \in \mathcal{G}_r(\theta)} \{\lambda_F F_w(Y, G) + \lambda_O O_w(Y, G)\}. \quad (1)$$

Both components use the same admissible comparison. When reporting them separately, use a measurable minimizer $G^\dagger(Y, r, \theta)$. If every effective weight $\lambda_F w_{rc}^F, \lambda_O w_{rc}^O$ is positive, loss is zero exactly when $Y \in \mathcal{G}_r(\theta)$; zero weights can enlarge the zero-loss set. Expectations use a prior on $\theta \mid r$, not observable Gold membership. All correspondences, policies, losses, and probability kernels are measurable. Expectations are finite unless an extended-real value represents infeasibility.

3.2 Resources and routing

Let x be the accumulated learning state and a resource intensity, which summarizes compute, information, and human or model time. Conditional on an implemented production and verification regime, expected loss is $\ell(r, a, x)$. For a per-task target $\bar{\ell}(r)$, define the resource requirement

$$z(r, x) = \inf\{a : \ell(r, a, x) \leq \bar{\ell}(r)\}. \quad (2)$$

An unattainable target has requirement $+\infty$. For a variable tolerance, write $z(r, x; \varepsilon) = \inf\{a : \ell(r, a, x) \leq \varepsilon\}$, so that $z(r, x) = z(r, x; \bar{\ell}(r))$.

The nested-resource benchmark has routine intensity $\underline{a}$ and high-level intensity $\bar{a} > \underline{a}$. An architecture installs finitely many intensities,

$$S = \{a_0, \dots, a_k\}, \quad a_0 = \underline{a} < a_1 < \dots < a_k = \bar{a}. \quad (3)$$

There are k intervals and $L_{\text{lay}} = k + 1$ active layers. Routine production P uses $\underline{a}$; high-level adjudication H uses $\bar{a}$ and remains fallible. Execution costs $c(a)$, a continuous, nonnegative, increasing function. There is no congestion. Architecture overhead $\mathcal{F}(S)$ covers installation, interfaces, calibration, and maintenance, excluding costs common to the two endpoints. Let $\mathcal{O}_t$ be the curator's information when routing date- t work.

Assumption 1 (Nested resources and observable requirements). For every task and learning state in the benchmark,

$$\ell(r, a, x) \leq \bar{\ell}(r) \iff a \geq z(r, x), \quad z(r, x) \in [\underline{a}, \bar{a}],$$

and $z(r, x)$ is $\mathcal{O}_t$ -measurable at routing. For some common $\underline{f} > 0$, every admissible architecture satisfies

$$\mathcal{F}(S) \geq \underline{f}(|S| - 2).$$

The first restriction makes technically sufficient resources nested and their minimum observable. Routing uses this assumed information about resource requirements, while correct records remain latent. The uniform overhead bound prevents unbounded optimal depth whenever an endpoint-only plan has finite cost. The exact-depth benchmark uses $\mathcal{F}(S) = f(k - 1)$, with $f > 0$.

The least costly feasible assignment is

$$a_S(z) = \min\{a \in S : a \geq z\}, \quad a_S(\underline{a}) = \underline{a}. \quad (4)$$

Layer a_j serves requirements in $(a_{j-1}, a_j]$. Dynamic optimality also requires the common-learning condition below. Appendix D relaxes observed requirements and allows routing to affect learning. With costless continuous customization, the present rationale for discrete layers would disappear.

3.3 Assignments, assessments, and compensation

Risk-neutral workers come from one population. In batch t , production and verification assignments are disjoint: $\mathcal{I}_t^P \cap \mathcal{I}_t^V = \emptyset$. Workers may change roles across batches. Each role involves costly private inputs and reporting choices, but no worker supplies both roles' effort in the same batch. No common effort budget is imposed.

Producer i chooses a complete strategy u_i , including information acquisition, computing, reporting, and nonresponse. The strategy costs $C_i^P(u_i)$ and commits a submission $\hat{Y}_i$, with expected loss

$$L_P(u_i; x) = \mathbb{E}[d_\lambda(\hat{Y}_i, \mathcal{G}_r(\theta)) \mid u_i, x].$$

Production assessment applies to that submission. The final release Y_i may incorporate subsequent repairs; its loss still includes required items that are never released.

The ordinary verification policy ν specifies assessment methods, routing, review assignments, aggregation, and recommended review strategies. It produces an assessment X_i from available references and current adjudication. The assessment rule is committed before production; the realized reference need not be. A producer's submission can trigger review, but its own review reports cannot improve its assessment. Any remaining influence enters the deviation comparisons.

Verifier j chooses a complete strategy v_j at cost $C_j^V(v_j)$, producing ordinary reports and responses to concealed synthetic controls. The protocol grades the latter through A_j , preserves access to ordinary sources, and commits responses before disclosing synthetic origin or feedback. Normalize $X_i, A_j \in [0, 1]$. Workers' errors may be correlated through shared information, models, difficult cases, or ownership.

Assessment is distinct from latent quality. For example, fix a normalization $\bar{L}_P > 0$ with $0 \leq L_P \leq \bar{L}_P$, and write $q_P = 1 - L_P/\bar{L}_P$. If a maintained measurement model gives

$$|\mathbb{E}[X_i \mid u_i; \nu, x] - q_P(u_i; x)| \leq \beta_P \quad \text{for every } u_i,$$

a quality gap above $2\beta_P$ implies a positive assessment gap. Neither agreement nor synthetic grades identify β_P . Appendix A separates validity, behavioral incentives, and identification of ordinary quality.

Payoffs and continuation. An aggregate reference Y_H is produced by n verifiers. Verifier j 's private input vector is $e_j \in \mathcal{E}_j \subseteq \mathbb{R}_+^{K_e}$, where K_e is the number of effort dimensions. Controlled automated execution creates no additional hidden-action choice, but consumes resources and may err.

At an information history h , let O collect contractible outcomes, $t_i^s(O)$ be the current transfer in role $s \in \{P, V\}$, h' the next history, and $V_i(h')$ the worker's net continuation value. The worker discount factor is $\delta_w \in (0, 1)$. Holding other workers' strategies and consistent beliefs fixed gives

$$\begin{aligned} U_i^P(u_i; h) &= \mathbb{E}[t_i^P(O) + \delta_w V_i(h') \mid u_i, h; \nu, x] - C_i^P(u_i; h), \\ U_j^V(v_j; h) &= \mathbb{E}[t_j^V(O) + \delta_w V_j(h') \mid v_j, h] - C_j^V(v_j; h). \end{aligned} \tag{5}$$

Continuation includes future production access, verification obligations, waiting, exit, and re-entry. These values must be generated by the announced policy.

For the scalar-effort and score-gradient arguments, let T_j be fixed compensation and $\mathbf{s}_j(e_j) \in \mathbb{R}^{K_{\text{score}}}$ sufficient statistics for its variable part, where K_{score} is their number. With assessment, settlement, and continuation fixed by ξ , write

$$U_j(e_j) = T_j + \Gamma_j(\mathbf{s}_j(e_j); \xi) - C_j(e_j). \tag{6}$$

Here C_j is the corresponding effort-cost specialization of C_j^V . Nonlinear settlement may require outcome probabilities, not only expected scores, in $\mathbf{s}_j$. Expected compensation need not equal compensation evaluated at expected scores. Participation requires $U_j(e_j^*) \geq \bar{U}_j$, where $\bar{U}_j$ is the outside option. Fixed transfers must respect funding and payment limits, including their effects on future opportunities.

3.4 Synthetic supervision

From a plausible, not necessarily certified record $\hat{y}$, the curator draws a mutation using $M_r(\cdot \mid \hat{y})$, records the change, and conceals the altered case among ordinary assignments. Its detection or removal can be graded without identifying the correct ordinary replacement. Let $\pi \in [0, 1]$ be the synthetic share and $w > 0$ a direct payment per pass. For scalar effort e , let $q(e)$ denote ordinary reliability and $p_A(e)$ the synthetic pass probability. The scalar benchmark is

$$p_A(e) = b_A + \vartheta_A q(e), \quad \vartheta_A = \rho\tau(1 - \zeta). \quad (7)$$

Here b_A is effort-independent success, and the loadings $\rho, \tau, 1 - \zeta \in [0, 1]$ represent negative-label validity, transfer of incentives to ordinary work, and concealment. The specification must give $p_A(e) \in [0, 1]$ throughout the effort domain.

A sufficient interpretation of the affine pass law puts an effort-independent weight ϑ_A on valid, relevant, concealed checks and makes expected scores on the remainder independent of effort. Sequential conditional loadings suffice; mutual independence is not required. Replacing ρ by a validity lower bound requires a justified incentive-gap bound. Lemma A.3 allows invalid controls to favor deviations instead. Zero loading or locally flat settlement supplies no local incentive through the synthetic score.

For m_H ordinary adjudication assignments, the additional synthetic volume is

$$N_A(m_H, \pi) = \frac{\pi}{1 - \pi} m_H. \quad (8)$$

Its cost $C_A(\pi, m_H)$ includes mutation, adjudication, administration, and capacity use. Direct payments satisfy $0 < w \leq \bar{w} < \infty$. When verification earns eligibility, calculate its incentive from the transition rule and the net continuation values generated by the policy.

3.5 Learning and implementable plans

A usable ordinary resolution can provide a reusable rule, training example, or retrieval method. Let Z_t be a recurring decision's resource requirement after earlier learning. Conditional on a usable resolution, impose monotone learning on the kernel P :

$$P([a, z] \mid z) = 1. \quad (9)$$

The requirement can fall without reaching routine intensity, but any move to cheaper resources must preserve the quality target. New task families may enter with high requirements. Rejected work still consumes resources, and unresolved demand remains in the workload. The learning kernel describes resolutions whose lessons can be reused at the maintained quality target. Operational acceptance uses available evidence.

A complete supervision policy $\mathcal{D}$ specifies synthetic exposure, panel size, mutations, assignments, assessments, eligibility transitions, and selection. Compensation policy Γ specifies transfers and future opportunities. Together with S , recommended strategies, and routing, these objects form a curation plan. Capacities, assignment rules, routing rules, assessment, and settlement are committed before workers act. Current reports may enter the committed assessment rule; learning changes only later model versions. Let $\mathcal{U}_i(h)$ and $\mathcal{V}_j(h)$

contain the complete feasible strategies for the two roles, and let $\bar{U}_i^P(h)$ and $\bar{U}_j^V(h)$ be their outside-option values.

Definition 1 (Implementable curation plan). A plan is implementable if its own outcome laws, consistent beliefs, and funded continuation values satisfy, at every relevant information history,

$$\begin{aligned} u_i^*(h) &\in \arg \max_{u \in \mathcal{U}_i(h)} U_i^P(u; h), & U_i^P(u_i^*; h) &\geq \bar{U}_i^P(h), \\ v_j^*(h) &\in \arg \max_{v \in \mathcal{V}_j(h)} U_j^V(v; h), & U_j^V(v_j^*; h) &\geq \bar{U}_j^V(h), \end{aligned}$$

and routing is sequentially optimal under those laws. The plan must also satisfy the information, quality, resource, and payment restrictions specified here.

All continuation values come from the same policy, as formalized in Appendix B. Reports, aggregation, and acceptance must generate the claimed usable-resolution supply, learning law, and workload. Payment, routing, and operational acceptance use available information; latent-loss restrictions evaluate the plan’s fidelity. Implementation is assessed over each role’s complete feasible strategy set.

Holding learning opportunities fixed across architectures. Fix a high-level discovery cohort and an implemented production regime. Let $\mathfrak{D}(S)$ be the set of supervision and compensation policies feasible with architecture S . For a fixed policy pair $(\mathcal{D}, \Gamma)$, let $Z_t^{S, \psi; \mathcal{D}, \Gamma}$ be a recurring decision’s potential requirement under feasible routing rule ψ . Let $R^{S, \psi}(\mathcal{D}, \Gamma)$ be the expected usable-resolution supply. Architectures are compared on a common realization of task arrivals and learning opportunities.

Assumption 2 (A common frontier of potential learning). The compared architectures share a feasible policy set $\mathfrak{D}$, so $\mathfrak{D}(S) = \mathfrak{D}$ for every compared S . For each $(\mathcal{D}, \Gamma) \in \mathfrak{D}$, there are a common supply $R(\mathcal{D}, \Gamma)$ and potential process $Z_t^{\mathcal{D}, \Gamma}$ such that

$$\begin{aligned} R^{S, \psi}(\mathcal{D}, \Gamma) &= R(\mathcal{D}, \Gamma), \\ Z_t^{S, \psi; \mathcal{D}, \Gamma} &= Z_t^{\mathcal{D}, \Gamma} \quad \text{almost surely for all } t, S, \psi. \end{aligned}$$

The benchmark holds discovery, production, and supervision feasibility fixed: architecture changes when learning is exploited and its cost, not the potential learning path. Eligibility incentives must therefore remain feasible for every compared architecture. Section 10 allows architecture to change learning or verification rents. When only π varies along a specified common supervision frontier, write $R(\pi)$ for the expected number of usable resolutions in the fixed cohort.

Let M_t be a nonrandom mass of recurring decisions and $\delta \in (0, 1]$ the workload discount factor, distinct from δ_w . Assume $\sum_t \delta^t M_t < \infty$. For a measurable set of requirements B_z , define the discounted workload measure

$$\mu_\pi(B_z) = \sum_t \delta^t M_t \mathbb{P}_\pi \{Z_t \in B_z\}. \quad (10)$$

This measure counts initial heterogeneity and every later visit, including unresolved demand. It values reuse through the decisions that benefit, not through a separate learning bonus. Under Assumption 2, least-cost feasible routing is also dynamically optimal. When the full policy changes, write $\mu_{\mathcal{D},\Gamma}$ and recompute the induced behavior, resolutions, transitions, workload, and routing.

3.6 Resource costs and the curator's choice

Each input is charged once. Let $C_{\text{sup}}(\mathcal{D}, \Gamma)$ collect discovery, verification effort, aggregation, and synthetic supervision not already included in execution cost $c(a)$. Production effort enters execution; verification effort enters supervision. Architecture overhead is measured over the workload horizon or as its stationary annuity. For a given fidelity target, the nested-resource problem is

$$\inf_{S, \mathcal{D}, \Gamma} \left\{ C_{\text{sup}}(\mathcal{D}, \Gamma) + \mathcal{F}(S) + \int c(a_S(z)) d\mu_{\mathcal{D}, \Gamma}(z) \right\}, \quad (11)$$

where the infimum is over implementable plans. A minimum is attained when the feasible set is nonempty and closed, the objective is lower semicontinuous, and the relevant choice sets are compact, including a bound on depth. Strict implementation without a fixed margin may yield only an infimum.

Transfers finance resources; they are not additional resource use. Let B_t^{pay} be the expected payment budget and $\bar{t}_{it}^s \geq 0$ individual payment caps. Feasible financing requires

$$\sum_{i \in \mathcal{I}_t^P} \mathbb{E}[t_{it}^P] + \sum_{j \in \mathcal{I}_t^V} \mathbb{E}[t_{jt}^V] \leq B_t^{\text{pay}}, \quad 0 \leq t_{it}^s \leq \bar{t}_{it}^s, \quad s \in \{P, V\}. \quad (12)$$

A realized payment-pool cap adds a statewise restriction. Future transfers are counted when paid, not again through their continuation values. Adding rewards to resource cost would double-count inputs unless a financing wedge were modeled.

Curator and worker accounts. Fix an external workload of N source records and their required releases. Submissions, accepted fields, and duplicate work do not change this denominator; omissions continue to incur loss. The curator's validation account $\mathcal{C}_N^v$ includes its computing, judgment, controls, assessment, aggregation, infrastructure, and training. The worker account $\mathcal{C}_N^m$ includes all production, review, synthetic responses, search, and training, whether or not rewarded. At common opportunity prices, with shared overhead allocated once,

$$\mathcal{C}_N^{\text{tot}} = \mathcal{C}_N^v + \mathcal{C}_N^m, \quad \bar{c}_N^s = \frac{\mathcal{C}_N^s}{N}, \quad s \in \{v, m, \text{tot}\}. \quad (13)$$

Worker review belongs in the worker account. Outsourcing can therefore lower the curator's bill without lowering total resource use. These accounts allocate execution and supervision costs between their bearers. Payments and fees finance activity, while their administration uses resources. At fixed prices and within the maintained technology class, Bronze costs $c_{\text{Br}} > 0$ for one routine execution.

Aggregate quality targets. Per-task fidelity targets can imply aggregate loss bounds. If only aggregate constraints are imposed, let a complete action ℓ , including release or abstention, have cost k_ℓ , expected inclusion loss f_ℓ , expected omission loss o_ℓ , and incremental continuation-cost saving $\Delta C_\ell^{\text{fut}}$. At a valid saddle point, routing minimizes

$$J_\ell(r, x) = k_\ell(r, x) + \Lambda_F f_\ell(r, x) + \Lambda_O o_\ell(r, x) - \delta \Delta C_\ell^{\text{fut}}(r, x), \quad (14)$$

where Λ_F, Λ_O are the associated loss shadow prices. Appendix D develops this comparison. Arbitrary common multipliers cannot rank nonconvex architectures; the depth theorem instead uses feasible per-task targets.

Table 1: Main notation

Symbol	Definition
r, P_t, θ	Observable task type, its date- t distribution, and latent state.
$\mathcal{C}_r, \hat{Y}_i, Y_i$	Candidate items, submitted record, and final release.
$\mathcal{G}_r(\theta), d_\lambda$	Admissible records and weighted source-fidelity loss.
$F_w, O_w; \lambda_F, \lambda_O$	Weighted inclusion and omission errors; their loss weights.
$x, \ell(r, a, x), \bar{\ell}(r)$	Learning state, expected loss, and the per-task loss target.
$a, z(r, x), c(a)$	Resource intensity, minimum required intensity, and execution cost.
$S, a_S(z), k, L_{\text{lay}}$	Installed intensities, routing, interval count, and layer count.
$\mathcal{F}(S), \underline{f}, \bar{f}$	Architecture overhead, its per-layer lower bound, and the uniform benchmark cost.
$\mathcal{I}_t^P, \mathcal{I}_t^V$	Disjoint production and verification assignment sets.
$u_i, v_j; C_i^P, C_j^V$	Complete strategies and their private costs in the two roles.
X_i, A_j, ν	Ordinary assessment, synthetic assessment, and ordinary review policy.
O, h, t_i^s, V_i	Contractible outcomes, information history, current transfer, and net continuation value.
δ_w, δ	Worker and workload discount factors.
$e_j, \mathcal{E}_j, K_e, \mathbf{s}_j$	Effort vector, feasible domain, dimension, and compensation statistics.
$\pi, w, p_A(e), q(e)$	Synthetic share, payment per pass, pass probability, and ordinary reliability.
$\rho, \tau, 1 - \zeta; \vartheta_A$	Validity, ordinary-work transfer, and concealment loadings; their product.
M_r, m_H, N_A, C_A	Mutation kernel, ordinary adjudication volume, synthetic volume, and its cost.
$R(\pi), Z_t, \mathbf{P}$	Usable-resolution supply, evolving requirement, and learning kernel.
$M_t, \mu_\pi, \mu_{\mathcal{D}, \Gamma}$	Decision mass and discounted workloads under restricted and full policies.

Table 1 continued

Symbol	Definition
$\mathcal{D}, \Gamma, C_{\text{sup}}$	Supervision policy, compensation policy, and supervision resource cost.
$N, \mathcal{C}_N^V, \mathcal{C}_N^m$	Fixed source workload and curator and worker resource accounts.
$\bar{c}_N^s, c_{\text{Br}}$	Resource cost per source record and the routine-execution baseline.

4 The incentive problem inside the reference

A review panel does not supply a trusted reference merely by being large. The following benchmark separates incentives to acquire information from independent and shared errors. Fixed pay and simple agreement provide a baseline for richer information-elicitation mechanisms (Miller et al., 2005; Dasgupta and Ghosh, 2013; Gao et al., 2019).

In a one-period comparison with no other payoff consequences, consider $\theta \in \{0, 1\}$ and an odd panel of size $n \geq 3$. Choosing $e_i = 0$ costs nothing and provides no information; $e_i = 1$ costs $c_{\text{eff}} > 0$ and produces a private signal. Workers can coordinate cheaply on the public focal report $z_{\text{foc}} \in \{0, 1\}$. The agreement payment is

$$T + b\mathbf{1}\{y_i = y_{k(i)}\}, \quad b > 0, \quad k(i) \neq i.$$

With truthful high effort, a common event makes all reports incorrect with probability $\kappa \in [0, 1)$. Outside that event, reports are conditionally independent and each is correct with probability $p \in (1/2, 1)$. Define

$$B_n(p) = \mathbb{P}\{\text{Bin}(n, p) \leq (n - 1)/2\}.$$

Producing each careful report uses $\chi > 0$ units of resources.

Fixed transfers may secure participation but do not induce costly acquisition. Agreement rewards, meanwhile, permit an equilibrium in which everyone submits the focal report. Even under truthful care, majority error equals $\kappa + (1 - \kappa)B_n(p)$. Convergence to the common-error floor requires $n\chi \rightarrow \infty$. The proof, a bound on error, and a sufficient panel size are given in Lemma A.1. Expertise can improve the signals that are acquired; aggregation neither acquires missing information nor removes errors shared by all reports.

5 Anti-Gold: from valid negatives to useful effort

A justified negative label is useful only if detecting it rewards behavior that improves ordinary work. The analysis connects the construction and cost of such controls to their incentive effects, first for a single effort margin and then for several. The final step identifies when the resulting effort produces better ordinary resolutions rather than merely higher synthetic scores.

5.1 Valid negatives without certified positive seeds

Fix the predicate being tested. Let $\mathcal{Y}_r$ be its finite response space and $\mathcal{G}_r^{\text{test}}(\theta)$ the values admissible in the given source context. For a test of the whole output, these objects are $2^{\mathcal{C}_r}$ and $\mathcal{G}_r(\theta)$. For a field-level test, admissibility concerns faithful values of that field. An error elsewhere in a record does not show that the particular mutation being tested is an error.

Definition 2 (Validity of a known negative). For $\epsilon_r \in [0, 1]$, a mutation kernel is ϵ_r -valid if

$$\sup_{\hat{y}} M_r(\mathcal{G}_r^{\text{test}}(\theta) \mid \hat{y}) \leq \epsilon_r \quad \text{for every maintained } \theta.$$

The bound $\rho_r = 1 - \epsilon_r$ concerns negative-label validity across draws; it does not reveal the realized label of each draw.

Suppose no more than m_r values are admissible and no mutation outcome has probability above $\bar{p}_r$. Then the chance of an accidentally valid outcome is bounded by $\min\{1, m_r \bar{p}_r\}$, as shown in Lemma A.2. With one admissible value and uniform sampling from N_r alternatives, the bound is $1/N_r$. Numerical fields, identifiers, offsets, dates, and relations can have large response spaces. Binary fields, synonymous expressions, and short lists need not. A demonstrable structural violation, such as a failed checksum, can yield $\epsilon_r = 0$. Evidence is required for both the admissible-support bound and the mutation distribution.

5.2 Comparing controls by incentive-adjusted cost

Fix the direct-payment technology in (7) and a common payment w . For this cost comparison only, positive-control search uses independent proposals and a perfect external certifier. A proposal costs g_0 , is admissible with probability $\alpha_{\text{adm}} \in (0, 1)$, and costs h_+ to certify. The relevant rarity is admissibility under this proposal distribution, not ordinary production's error rate. Repeated positive search and direct negative construction cost

$$c_+ = \frac{g_0 + h_+}{\alpha_{\text{adm}}} + k_+, \quad c_A = g_0 + h_A + k_A. \quad (15)$$

Here h_A allocates the cost of validating a negative family, while k_+, k_A cover execution and administration. The proposal is invalid with probability $1 - \alpha_{\text{adm}}$. If positive certification is strategic, its incentive costs and errors must also be included.

For positive controls, use pass law $b_+ + \vartheta_+ q(e)$; for negative controls, use $\vartheta_A = (1 - \alpha_{\text{adm}})\tau(1 - \zeta)$. Both loadings are positive, and both pass laws take values in $[0, 1]$. Exposure levels v_A, v_+ provide equal marginal incentives when

$$v_A \vartheta_A = v_+ \vartheta_+. \quad (16)$$

After imposing reward, exposure, participation, and payment limits, Anti-Gold provides the covered scalar incentive more cheaply when

$$\frac{g_0 + h_A + k_A}{(1 - \alpha_{\text{adm}})\tau(1 - \zeta)} < \frac{(g_0 + h_+)/\alpha_{\text{adm}} + k_+}{\vartheta_+}. \quad (17)$$

Effort-independent differences in expected compensation can be offset with fixed transfers only to the extent permitted by those limits. Rarity favors negative controls when their cost remains bounded and their transport and concealment loadings remain positive.

A visibly artificial defect may be detectable without effort useful for ordinary work. Likewise, a reward for rejection alone may make rejecting everything optimal. Selective detection or correction must therefore be supported by rewards for legitimate preservation, ordinary controls, or complementary scores. Negative checks alone establish neither a correct replacement nor a complete record.

If the probability of an invalid control is at most ϵ and the high-effort pass advantage on valid controls is at least Δp_{val} , the overall advantage can be as low as $(1 - \epsilon)\Delta p_{\text{val}} - \epsilon$ (Lemma A.3). Even high validity can therefore leave a negative incentive gap. A product lower bound requires invalid controls not to favor deviations. The affine model further assumes effort-independent residual scores. Covering a positive incentive deficit requires a justified positive pass gap.

5.3 Displacing low-effort coordination

Unless reoptimization is specified, hold ordinary compensation, peers, beliefs, and settlement fixed. Informative behavior H costs $c_{\text{eff}} > 0$; focal behavior L costs zero. Their ordinary compensation levels are B_H, B_L . Assume

$$\Delta p_A = \rho\tau(1 - \zeta)(q_H - q_L) > 0, \quad K = c_{\text{eff}} + B_L - B_H.$$

A check arrives with probability $\pi \in [0, \bar{\pi}]$ and pays $w > 0$ for a pass. Consider perfectly substitutable exposures $v_A, v_R \geq 0$ to covered controls, with positive cost-loading pairs $(c_A, \iota_A), (c_R, \iota_R)$. To obtain utility margin $\eta \geq 0$, the exposures must satisfy

$$\iota_A v_A + \iota_R v_R \geq \Delta_\eta, \quad \Delta_\eta = [K + \eta]_+. \quad (18)$$

For the least-cost formula, exposures are uncapped; caps are considered below. The ordinary-resolution control remains fallible and needs its own justification. The case $\eta = 0$ requires weak implementation, while $\eta > 0$ fixes a positive margin. The incentive comparison is between H and L , holding peers' behavior fixed.

Proposition 1 (Implementing effort with least-cost covered controls).

$$H \succeq L \iff w\pi\Delta p_A \geq K, \quad H \succ L \iff w\pi\Delta p_A > K. \quad (19)$$

If $K > 0$, strict implementation is feasible if and only if $\pi^{IC} < \bar{\pi}$, with

$$\pi^{IC} = \frac{K}{w\rho\tau(1 - \zeta)(q_H - q_L)} \quad (20)$$

the infimum of strictly implementing exposures whenever that set is nonempty. The minimum cost of satisfying (18) is

$$C_\eta^* = \Delta_\eta \min \left\{ \frac{c_A}{\iota_A}, \frac{c_R}{\iota_R} \right\}. \quad (21)$$

Proof of Proposition 1. The payoffs to high effort and the focal strategy are, respectively, $T + B_H + w\pi p_A^H - c_{\text{eff}}$ and $T + B_L + w\pi p_A^L$. Subtracting gives $w\pi\Delta p_A - K$, which proves (19). For $K > 0$, the strict version of this inequality yields (20); some admissible exposure satisfies it precisely when $\pi^{IC} < \bar{\pi}$.

A utility gain of at least η is equivalent to $\iota_A v_A + \iota_R v_R - K \geq \eta$. Nonnegative exposures and loadings turn this into (18). Control j supplies a unit of incentive at cost c_j/ι_j . If a more expensive source is used while a cheaper source is uncapped, replacing its incentive contribution with the cheaper source lowers total cost. This exchange proves (21) and the allocation rule. Equal cost-loading ratios leave all binding mixtures equally costly. The same exchange establishes the merit order under caps, provided available capacity covers Δ_η . For $\Delta_\eta = 0$, the cost minimum is attained with no exposure. $\square$

When $\Delta_\eta > 0$, Anti-Gold alone is optimal under

$$\frac{c_A}{\iota_A} < \frac{c_R}{\iota_R}. \quad (22)$$

The reversed inequality selects ordinary controls alone; equality makes every binding mixture optimal. If $\Delta_\eta = 0$, choose zero exposure. With $v_j \leq \bar{v}_j$, feasibility requires $\iota_A \bar{v}_A + \iota_R \bar{v}_R \geq \Delta_\eta$, and the cheaper incentive capacity is used first.

For $K < 0$, care is already strictly preferred without checks. At $K = 0$, no checks leave the worker indifferent, while any positive exposure makes care strictly preferable. If $K > 0, w \leq \bar{w}$, an exposure-reward pair strictly implements care exactly when $K < \bar{w}\bar{\pi}\Delta p_A$. For $K \geq 0$, strict implementation without a fixed positive margin or a favorable tie-break may admit an infimum but no minimizing policy.

When ordinary compensation is paid exclusively on ordinary assignments, use $B_j = (1 - \pi)b_j$. If $\Delta b_{\text{ord}} = b_L - b_H \geq 0$, the exposure infimum becomes $(c_{\text{eff}} + \Delta b_{\text{ord}})/(w\Delta p_A + \Delta b_{\text{ord}})$, rather than (20).

5.4 Continuous effort and current incentive leverage

Hold ordinary pay fixed and let effort vary over a convex scalar interval $\mathcal{E}$. Assume diminishing returns to effort and strictly convex effort costs: q, C are twice differentiable, with $q' > 0, q'' \leq 0, C' > 0, C'' > 0$. Maintain $w\vartheta_A > 0$ and utility

$$U(e) = T + w\pi\{b_A + \vartheta_A q(e)\} - C(e).$$

Strict concavity makes $w\pi\vartheta_A q'(e) = C'(e)$ the condition for an interior optimum. At an included lower endpoint the one-sided derivative is nonpositive; at an included upper endpoint it is nonnegative. An interior target $\hat{e}$ is implemented if and only if

$$\pi(\hat{e}) = \frac{C'(\hat{e})}{w\vartheta_A q'(\hat{e})} \in (0, \bar{\pi}]. \quad (23)$$

Along a differentiable interior branch, implicit differentiation yields

$$\frac{de^*(\pi)}{d\pi} = \frac{w\vartheta_A q'(e^*)}{C''(e^*) - w\pi\vartheta_A q''(e^*)} > 0. \quad (24)$$

Consequently, one check supports an expected number of ordinary assignments equal to

$$\mathcal{L}(\hat{e}) = \frac{1 - \pi(\hat{e})}{\pi(\hat{e})}. \quad (25)$$

This ratio measures leverage over current effort. An ordinary resolution's future reuse value depends on the later workload it benefits.

5.5 Which effort dimensions are rewarded?

Fix assessment, settlement, and reporting continuation. On a convex effort domain, assume differentiable, strictly concave utility and consider an interior target. Define $\mathcal{J}_j = D_{e_j} \mathbf{s}_j$ and $\Omega_j = \nabla_{\mathbf{s}} \Gamma_j$. The target is implemented exactly when $\mathcal{J}_j^\top \Omega_j = \nabla C_j$. Unrestricted affine weights can generate $\text{range}(\mathcal{J}_j^\top)$; nonnegative weights generate the score-gradient cone, and bounded admissible weights generate only their image (Lemma A.4). These gradient tests give global incentives under the maintained strict concavity, with payments separately subject to feasibility.

A detection score may leave omission search, evidence recovery, or supported additions without rewards. Content that appears in neither the candidate set nor assigned searches lies outside the observed universe. Source searches and omission audits must create that coverage. Within each role, the relevant comparisons include information acquisition, reporting, and nonresponse. Repeating checks can improve precision without changing the score-gradient span; it cannot reward a direction to which every score is insensitive.

Example 1 (Detection incentives leave omission search uncovered). A verifier chooses detection effort e_D and omission-search effort e_O over $[0, 1]^2$, incurring $C(e_D, e_O) = (e_D^2 + e_O^2)/2$. A normalized detection score $A_D \in [0, 1]$ has mean $\bar{A}_D(e_D, e_O) = e_D$, maintaining the validity, transport, and reporting conditions needed for that relationship. Under payment $T + b_D A_D$,

$$U = T + b_D e_D - \frac{1}{2}(e_D^2 + e_O^2), \quad e_D^* = \min\{b_D, 1\}, \quad e_O^* = 0 \quad (b_D \geq 0).$$

Neither a higher detection reward nor additional repetitions of the same test implement positive omission search. The gradient $(1, 0)$ contains no component in that direction.

Now add a separately justified, source-grounded discovery or omission-control score $A_O \in [0, 1]$, with mean $\bar{A}_O = e_O$. The contract $T + 0.6A_D + 0.4A_O$ uniquely induces $(e_D, e_O) = (0.6, 0.4)$, since strict concavity makes both first-order conditions globally sufficient. With $T = 0$, expected compensation is 0.52, effort costs 0.26, and net utility is 0.26. Thus the zero outside option is satisfied, and the largest possible payment is one. The omission-control score requires its own source-grounded validity and grading justification. Removing content from an uncertified record does not automatically produce a valid omission test.

5.6 From ordinary repair to reusable resolutions

Hold production behavior and the ordinary cohort fixed. Use the interior scalar-effort response above, differentiable repair rates, and a fixed source-aligned admissible comparison. Let $F_0, G_0, O_0 \geq 0$ denote disjoint weighted amounts of false, correct, and omitted content. Write

r_1, r_2 for the fractions repaired, and r_0^F, r_0^O for newly introduced errors per unit of initially correct content. Loss is

$$L_H(e) = \lambda_F [F_0\{1 - r_1(e)\} + G_0 r_0^F(e)] + \lambda_O [O_0\{1 - r_2(e)\} + G_0 r_0^O(e)]. \quad (26)$$

No normalization $F_0 + G_0 = 1$ is imposed. If errors interact or the closest admissible comparison changes, the appropriate definition is (1) instead.

Along an interior effort response,

$$\frac{dL_H}{d\pi} = -\frac{de^*}{d\pi} [\lambda_F \{F_0 r_1' - G_0 (r_0^F)'\} + \lambda_O \{O_0 r_2' - G_0 (r_0^O)'\}].$$

Because $de^*/d\pi > 0$, ordinary loss falls locally if and only if

$$\lambda_F \{F_0 r_1'(e^*) - G_0 (r_0^F)'(e^*)\} + \lambda_O \{O_0 r_2'(e^*) - G_0 (r_0^O)'(e^*)\} > 0. \quad (27)$$

For the fixed cohort, $R'(\pi) > 0$ follows if the probability that a resolution is usable is differentiable in L_H with strictly negative derivative.

When production incentives or review queues change, recompute R using the resulting behavior and workload. With eligibility rewards, derive the effort response from the policy's continuation values.

6 Implementing production and verification separately

Verification must be induced before its assessments can support production incentives. The two roles therefore require separate constraints within one funded policy. Direct payments and future access can both reward review; a two-state policy shows how production income determines the value of that access. Throughout, fix the relevant history, other workers' strategies, and consistent beliefs, and use continuation values generated by the policy being evaluated.

6.1 The production constraint

Keep the ordinary assessment rule committed, but allow routing and judgments to respond to each deviation under that rule. Let $\mathcal{U}_i(h)$ contain complete feasible strategies, including reporting and nonresponse, and define

$$G_i^P(u; h) = \mathbb{E}[t_i^P(O) + \delta_w V_i(h') \mid u, h; \nu, x].$$

Implementing the target requires

$$G_i^P(u_i^*; h) - G_i^P(u; h) \geq C_i^P(u_i^*; h) - C_i^P(u; h) \quad \text{for every } u \in \mathcal{U}_i(h), \quad (28)$$

as well as participation relative to declining the assignment or exiting.

If a cheaper strategy induces the same complete payment-relevant law, it is strictly preferred. This includes copying a fallible reference whenever it reproduces the target's payment-relevant law at lower cost. Lemma B.1 extends the comparison to mixtures with bounded compensation. Synthetic grades influence production incentives only through the ordinary assessment that they help induce.

Rank-based production pay. When compensation is a prize, G_i^P depends on payable-rank probabilities, not mean assessment scores (Lazear and Rosen, 1981). A scoring rule that elicits truthful reports under direct payment need not do so when only the highest score wins (Witkowski et al., 2023). The full constraint (28) accommodates such rewards by evaluating each deviation’s actual payment and continuation law. Changing entry or prizes also changes the future rents used to reward verification.

6.2 The verification constraint: pay now or access later

For the complete strategy set $\mathcal{V}_j(h)$, write $G_j^V(v; h) = \mathbb{E}[t_j^V(O) + \delta_w V_j(h') \mid v, h]$. Verification is incentive-compatible when

$$G_j^V(v_j^*; h) - G_j^V(v; h) \geq C_j^V(v_j^*; h) - C_j^V(v; h) \quad \text{for every } v \in \mathcal{V}_j(h), \quad (29)$$

and participation also holds. Under its behavioral assumptions, Proposition 1 supplies the direct-payment component.

6.3 Binary incentives in the two roles

Compare careful behavior H with a specified alternative L in each role. For production, care lowers submitted loss and costs an additional $c_P > 0$. Take $X_i \in \{0, 1\}$, let $p_e(\nu, x) = \mathbb{P}(X_i = 1 \mid e; \nu, x)$, use compensation $T_P + B_P X_i$ with $B_P > 0$, and hold production continuation independent of current behavior. For verification, let $c_V > 0$ be the extra cost of care and ΔT_V its current-payment advantage. Careful and careless verification lead to eligibility with probabilities g_H, g_L , respectively. The state values V_E, V_I must summarize all strategy-dependent continuation effects: conditional on either state, value cannot also depend on the current regime.

The probabilities incorporate the score history, exposure, and updating rule. Fallible assessments can support care when admissible rewards distinguish it from shortcuts enough to cover its additional cost.

Proposition 2 (Incentive support for the two assigned roles). *Both roles weakly prefer H to L if and only if*

$$B_P[p_H(\nu, x) - p_L(\nu, x)] \geq c_P, \quad (30)$$

$$\Delta T_V + \delta_w(g_H - g_L)(V_E - V_I) \geq c_V. \quad (31)$$

Strict preference in either role replaces its inequality by a strict inequality.

Proof. Subtracting careless production utility from careful production utility cancels the fixed payment and the common continuation value. The remaining difference is the left side of (30) minus c_P . For verification regime e , expected continuation is $g_e V_E + (1 - g_e) V_I$. Taking the same utility difference gives the left side of (31) minus c_V . Risk neutrality makes the payoff to an ex ante mixture the corresponding weighted average. Participation, unlike these incentive comparisons, evaluates target utility against the outside option. $\square$

If the two regimes exhaust a role's strategies, its inequality characterizes conditional best responses, including ex ante mixtures. The result allows correlated errors. Additional reporting, identity, or history-dependent deviations enter the full inequalities; participation is checked against the outside option.

With $\Delta T_V = 0$, future access can provide the entire incentive for care. The relevant value is expected income net of production costs, future verification, waiting, and re-entry, not the advertised prize. In particular, V_I can be positive. Inexpensive re-entry can neutralize exclusion, and a final costly review needs a terminal reward or another payoff consequence. A rise in mean score need not increase $g_H - g_L$ when eligibility uses a rolling threshold. More observations may reduce uncertainty without increasing incentives, while delayed feedback leaves privately informed contingent strategies available.

6.4 A two-state policy for eligibility

The continuation gap can be solved explicitly when workers alternate between production and verification. This reduced policy makes the dependence of review incentives on production income explicit.

Example 2 (Eligibility rents generated by a funded policy). An eligible worker E performs production, receives $T_P + B_P X$, and moves to I . A worker in I verifies without current pay, returning to E with probability g_e and otherwise remaining in I . Maintain $g_H > g_L \geq 0$, $g_H \leq 1$, $p_H > p_L$, and the two exhaustive strategies in each role. Care costs $c_P > 0$ in production and $c_V > 0$ in verification; the alternatives are costless. Exit is worth zero, and re-entry cannot restore eligibility. Assignments are stationary and feasible, with disjoint current roles.

Careful production earns net current income $r_P = T_P + B_P p_H - c_P$. The policy implies

$$V_E = r_P + \delta_w V_I, \quad V_I = -c_V + \delta_w \{g_H V_E + (1 - g_H) V_I\},$$

and hence

$$\Delta V_{\text{elig}} := V_E - V_I = \frac{r_P + c_V}{1 + \delta_w g_H}, \quad V_I = \frac{-c_V + \delta_w g_H \Delta V_{\text{elig}}}{1 - \delta_w}. \quad (32)$$

Substitution into the two incentive constraints gives

$$B_P(p_H - p_L) \geq c_P, \quad r_P \geq \frac{c_V(1 + \delta_w g_L)}{\delta_w(g_H - g_L)}. \quad (33)$$

The verification inequality also gives $V_I \geq 0$ and $V_E > 0$, ensuring participation. Strict inequalities give strict preference for care.

Production and verification need not fail at the same reward. With $T_P = 0$, the two constraints can be written as

$$B_P \geq \frac{c_P}{p_H - p_L}, \quad B_P \geq \frac{1}{p_H} \left[c_P + \frac{c_V(1 + \delta_w g_L)}{\delta_w(g_H - g_L)} \right].$$

When the second threshold is larger, a reward between them covers production effort under the intended review policy but fails to induce that review. These thresholds evaluate the

intended review policy. Below its verification threshold, equilibrium continuation values must be recomputed using the resulting review behavior.

Funding requires $T_P \geq 0$, $T_P + B_P \leq \bar{t}^P$, and a pool sufficient to pay assigned producers. The stationary production share is $g_H/(1 + g_H)$; all other workers verify. Capacity and the maintained assessment, validity, and behavioral-coverage conditions must support these assignments. Review resources are charged even when current review is unpaid.

6.5 When the recursive incentive problem is closed

The recursion is closed only if synthetic incentives induce sufficiently broad, selective ordinary verification to support feasible production incentives. The relevant deviations include blanket rejection and focal agreement. Current eligibility incentives must be backed by the future production opportunities used to calculate their value. Changing prizes, assignments, queues, or architecture therefore calls for renewed checks of both roles' constraints and of continuation consistency. Appendix B states the full-history requirement.

7 How reusable judgment lowers resource requirements

Usable ordinary resolutions can lower two resource requirements: the review needed to sustain production incentives and the execution needed to meet a quality target. The analysis links these savings to the propagation of learned capability, then considers feedback through future work opportunities. Resolutions may come from winning or losing entries. Planted defects can separately be retained as labeled training controls.

7.1 Propagation of capability

Fix architecture S , component gains $x_t \in \mathbb{R}_+^{L_{\text{lay}}}$, the usable-resolution input, and propagation coefficients. Write spr for spectral radius:

$$x_{t+1} = \mathbf{A}_S x_t + \mathbf{b}_S R(\pi), \quad \mathbf{A}_S \geq 0, \quad \mathbf{b}_S \geq 0, \quad \text{spr}(\mathbf{A}_S) < 1. \quad (34)$$

The system has the unique stable state $x^* = (I - \mathbf{A}_S)^{-1} \mathbf{b}_S R(\pi)$. Writing $\eta_j^{\text{learn}} = [(I - \mathbf{A}_S)^{-1} \mathbf{b}_S]_j \geq 0$ gives $dx_j^*/dR = \eta_j^{\text{learn}}$. Lemma D.1 derives the path expansion. When $R'(\pi) > 0$, a component improves strictly if and only if a positive path connects the input to that component. These capability gains have resource value when they change feasible assessment or execution.

7.2 Reducing adjudication at unchanged production incentives

Within a specified policy family, let x_J measure learned judging capability and d the intensity of costly adjudication. Hold fixed the production target H , alternative L , extra cost $c_P > 0$, and announced reward policy. Care's compensation advantage is

$$\mathcal{I}_P(d, x_J) = G^P(H; d, x_J) - G^P(L; d, x_J).$$

With a binary success payment, it is $B_P[p_H(d, x_J) - p_L(d, x_J)]$. A tournament instead requires the relevant payable-rank probabilities and continuation effects.

Across the capability range, suppose the same compact interval $[d_0, \bar{d}]$ of d is available. At recommended production, every policy in that interval satisfies verification incentives, participation for both roles, funding, and ordinary quality. Assume L is the sole remaining production deviation. Each complete policy generates its own continuation values. Let $k_V(d)$ be nondecreasing and include all verification resources for a fixed workload: automated assessment, strategic review, controls, and administration. Suppose $\mathcal{I}_P$ is continuous, nondecreasing in d, x_J , and satisfies $\mathcal{I}_P(\bar{d}, x_J) \geq c_P$. Define

$$d^*(x_J) = \min\{d \in [d_0, \bar{d}] : \mathcal{I}_P(d, x_J) \geq c_P\}. \quad (35)$$

Proposition 3 (Learning and the cost of implemented verification). *Both $d^*(x_J)$ and $k_V(d^*(x_J))$ are weakly decreasing in x_J . Along a differentiable interior branch where production incentives bind and $\partial_d \mathcal{I}_P > 0$,*

$$\frac{dd^*(x_J)}{dx_J} = -\frac{\partial_{x_J} \mathcal{I}_P(d^*, x_J)}{\partial_d \mathcal{I}_P(d^*, x_J)} \leq 0. \quad (36)$$

Proof. The upper-end condition and continuity ensure that the feasible set defining minimum review is nonempty and closed. Its minimum therefore exists. A larger x_J weakly raises $\mathcal{I}_P$ at each d , leaving every previously feasible intensity feasible. The minimum feasible intensity cannot increase. Nondecreasing k_V then implies the cost comparison. Differentiating the binding condition $\mathcal{I}_P(d^*(x_J), x_J) = c_P$ proves (36). Finally, combine the chain rule with $dx_J^*/dR = \eta_J^{\text{learn}}$ to obtain (37). $\square$

Where the derivatives exist, learning and adjudication savings combine as

$$\frac{dk_V(d^*(x_J^*(R)))}{dR} = -k'_V(d^*) \frac{\partial_{x_J} \mathcal{I}_P}{\partial_d \mathcal{I}_P} \eta_J^{\text{learn}} \leq 0. \quad (37)$$

The decrease is strict if $k'_V > 0$, $\partial_{x_J} \mathcal{I}_P > 0$, and $\eta_J^{\text{learn}} > 0$. A direct cost effect of x_J can be included when it is weakly decreasing.

With several deviations or policies that are not ordered by one intensity, the needed condition is stronger: every previously feasible policy must have a no-more-costly replacement preserving *all* incentive, participation, funding, and quality restrictions. This condition includes preserving feasible eligibility incentives as work opportunities change.

7.3 Execution savings and feedback through future opportunities

Fix the fidelity target $\bar{Q}$ and an implementable policy family. Let x_{Br} denote learned routine capability and $k_{\text{Br}}(x_{\text{Br}}; \bar{Q})$ its differentiable execution-cost function. Assume $\partial k_{\text{Br}} / \partial x_{\text{Br}} \leq 0$. Then

$$\frac{dk_{\text{Br}}(x_{\text{Br}}^*(R); \bar{Q})}{dR} = \frac{\partial k_{\text{Br}}}{\partial x_{\text{Br}}} \eta_{\text{Br}}^{\text{learn}} \leq 0. \quad (38)$$

Strict savings need both technological improvement and a positive learning path. Assessment gains and execution gains require support from held-out ordinary cases.

At unchanged effort, lower compensation relaxes the financing requirement, not resource use. Assessment and execution savings belong once in supervision or execution costs, rather than being counted again as a learning bonus. Lower prizes, more dilution, or longer waits may reduce $V_E - V_I$ below the level needed for verification. Both roles' incentive constraints must therefore be recomputed after such policy changes.

The linear propagation system takes R as fixed. When capability, production, or queue composition affects resolution supply, solve jointly for capability and resolutions. Proposition D.1 supplies a sufficient stability bound for this feedback.

8 Designing the intermediate resource layers

Intermediate layers save resources when enough current and future work can use them. A local insertion test measures those savings with existing cutoffs held fixed; global design also relocates the cutoffs. A uniform learning path then gives an exact depth rule and shows how supervision and layer choice interact. Throughout this section, maintain Assumptions 1 and 2: production and verification are implementable, and architectures exploit the same potential learning paths. Workload has finite expected discounted mass, and execution cost is continuous and increasing.

8.1 Overprovisioning and the value of one insertion

Fix supervision and the induced workload μ_π . Define overprovisioning by subtracting hypothetical continuously tailored execution from the cost of the installed architecture:

$$\mathcal{E}(S, \pi) = \int [c(a_S(z)) - c(z)] d\mu_\pi(z), \quad (39)$$

Continuous tailoring is used only as an accounting reference; it is not an available production technology. Inserting $a_{\text{new}} \in (a_-, a_+)$ reassigns precisely the requirements in $(a_-, a_{\text{new}}]$ from intensity a_+ to a_{new} . The insertion is beneficial if and only if $[c(a_+) - c(a_{\text{new}})]\mu_\pi((a_-, a_{\text{new}}]) > f_{\text{new}}$; equality gives indifference (Lemma D.2). All later visits generated by learning are already included in the occupation measure, so an additional reuse benefit would count them twice.

8.2 Choosing the whole partition

Fix supervision, charge $f > 0$ per intermediate layer, and break ties in favor of fewer layers. For the scaled-requirement comparison only, take affine costs $c(a) = c_0 + \gamma a$, $\gamma > 0$, and write $z = \underline{a} + \Delta_a \tilde{z}$, $\Delta_a > 0$, holding the normalized workload and learning law on $\tilde{z} \in [0, 1]$ fixed. For $a_- \leq a_+$, let

$$H_\pi(a_-, a_+) = \int_{(a_-, a_+]} [c(a_+) - c(z)] d\mu_\pi(z).$$

Denote by $E_j(a_+)$ the minimum overprovisioning on $(\underline{a}, a_+]$ with j intervals, allowing intervals of zero width. Initialize $E_0(\underline{a}) = 0$ and $E_0(a_+) = +\infty$ when $a_+ > \underline{a}$. Conditioning on the cutoff immediately before the last one yields

$$E_j(a_+) = \inf_{a_- \leq a_+} \{E_{j-1}(a_-) + H_\pi(a_-, a_+)\}. \quad (40)$$

Write $E_k = E_k(\bar{a})$. On a finite grid, the recursion is directly computable. Allowing coincident cutoffs compactifies the continuous problem; positive overhead then removes redundant layers. With continuous c , finite workload, and lower semicontinuity, minima exist even if interior requirements have atoms.

Proposition 4 (Globally optimal partitions and scaled requirements). *The optimal partition satisfies*

$$k^*(\pi) \in \arg \min_{k \geq 1} \{E_k + f(k-1)\}, \quad L_{\text{lay}}^*(\pi) = k^*(\pi) + 1, \quad (41)$$

Only counts with $f(k-1) \leq E_1$ require consideration, so the search is finite. The optimized overprovisioning costs satisfy

$$2E_k \leq E_{k-1} + E_{k+1}, \quad k \geq 2.$$

It follows that the smallest minimizing interval count is

$$k^*(\pi) = \min\{k \geq 1 : E_k - E_{k+1} \leq f\}. \quad (42)$$

Selected depth is weakly decreasing in f . In the scaled-requirement comparison, $E_k = \gamma \Delta_a \mathcal{E}_k^{\text{norm}}$, where $\mathcal{E}_k^{\text{norm}}$ is nonincreasing and does not depend on Δ_a . Thus selected depth is weakly increasing in $\gamma \Delta_a / f$.

The proof is in Appendix D.4.

This comparative static applies to scale changes in a fixed normalized distribution. The lower endpoint can shift to hold the mean fixed provided feasibility and affine costs are preserved. Moving mass from the midpoint to already installed endpoints can increase variance while removing all demand for a bridge.

Diminishing marginal gains apply to *reoptimized* partitions. For a unit uniform workload on $[0, 1]$ with $c(a) = a$, the optimal single bridge is at $1/2$. Keeping that bridge fixed while adding another leaves excess cost $3/16$; relocating both to $1/3, 2/3$ gives $E_3 = 1/6$. When $1/16 < f < 1/12$, a greedy insertion procedure stops at two intervals even though three are globally optimal. The stopping criterion must therefore use values from the recursion, not the gains from a sequence of frozen-cutoff insertions.

8.3 A uniform path from high-level to routine work

Maintain affine costs $c(a) = c_0 + \gamma a$, $\gamma > 0$, overhead $f > 0$ per intermediate layer, and the tie-break toward fewer layers. Let $\Delta_a = \bar{a} - \underline{a} > 0$, and distribute initial interior mass M_0 uniformly. Endpoint atoms are permitted and generate no overprovisioning. A high-level cohort contains N_{fam} potential resolution families, of which an expected $R(\pi) \in [0, N_{\text{fam}}]$ yield usable resolutions. Each family recurs at rate λ_{rec} decisions per unit of time. Over an undiscounted cycle of length T_{learn} , usable families follow the requirement path

$$z(t) = \bar{a} - \Delta_a \frac{t}{T_{\text{learn}}}, \quad 0 \leq t \leq T_{\text{learn}}. \quad (43)$$

Unresolved families remain at $\bar{a}$. Layers exploit this progress without changing its speed, and subsequent routine production is common endpoint workload. Setting $D = \lambda_{\text{rec}} T_{\text{learn}}$, the mass on an interior measurable set A is

$$\begin{aligned}\mu_\pi(A) &= \frac{M_0}{\Delta_a} |A| + R(\pi) \lambda_{\text{rec}} \int_0^{T_{\text{learn}}} \mathbf{1}\{z(t) \in A\} dt \\ &= \frac{\mathcal{M}(\pi)}{\Delta_a} |A|, \quad \mathcal{M}(\pi) = M_0 + DR(\pi).\end{aligned}\tag{44}$$

The unresolved mass $D[N_{\text{fam}} - R(\pi)]$ remains at the high endpoint and must still be counted in total cost.

Define

$$\Xi(\pi) = \frac{\gamma \Delta_a}{2} [M_0 + DR(\pi)].\tag{45}$$

Let $C_{\text{common}}(\pi)$ collect supervision and the ideal execution account $\int c(z) d\mu_\pi(z)$, including unresolved families at the high endpoint. These costs do not depend on installed depth. Locations are chosen conditional on depth before depths are compared.

Theorem 1 (Exact depth along a uniform routinization path). *The equally spaced capacities $a_j = \underline{a} + j\Delta_a/k$, $j = 0, \dots, k$, minimize cost for each fixed $k \geq 1$. The minimized cost is*

$$C_k(\pi) = C_{\text{common}}(\pi) + \frac{\Xi(\pi)}{k} + f(k-1).\tag{46}$$

The selected depth is

$$k^*(\pi) = \max \left\{ 1, \left\lceil \frac{\sqrt{1 + 4\Xi(\pi)/f} - 1}{2} \right\rceil \right\}, \quad L_{\text{lay}}^*(\pi) = k^*(\pi) + 1.\tag{47}$$

Proof of Theorem 1. Let $\Delta a_j = a_j - a_{j-1}$ be the interval widths, so $\sum_{j=1}^k \Delta a_j = \Delta_a$. Endpoint atoms require no excess intensity. Integrating the affine cost difference over the interior intervals gives

$$\mathcal{E}(S, \pi) = \frac{\gamma \mathcal{M}(\pi)}{2\Delta_a} \sum_{j=1}^k (\Delta a_j)^2.$$

By $\sum_j (\Delta a_j)^2 \geq \Delta_a^2/k$, equal widths minimize this expression. Taking $\Delta a_j = \Delta_a/k$ yields $\mathcal{E} = \Xi(\pi)/k$ and therefore (46). The difference between adjacent costs is $C_{k+1} - C_k = f - \Xi(\pi)/[k(k+1)]$, which is nondecreasing in k . The smallest minimizing count is the first integer with a nonnegative adjacent difference. Solving that condition gives (47), including $\Xi = 0$.

If $\Xi/f = j(j+1)$ for an integer $j \geq 1$, then $C_j = C_{j+1}$. Because $\Xi > 0$ at such a threshold, strict increase of the adjacent difference rules out any other minimizer, and the tie-break chooses j . For $\Xi = 0$, each increment is $f > 0$, so $k = 1$ is uniquely optimal. Finally, the selected integer is nondecreasing in Ξ/f , which gives the comparative static. $\square$

Equivalently, choose the least $k \geq 1$ satisfying $\Xi/[k(k+1)] \leq f$. An additional interval is worthwhile exactly when

$$\frac{\Xi(\pi)}{k(k+1)} > f. \quad (48)$$

At $\Xi/f = j(j+1)$, only the adjacent counts $j, j+1$ tie, and j is retained. Depth is weakly increasing in Ξ/f , and therefore in $\Delta_a, \gamma, M_0, DR(\pi)$ with the other primitives fixed, while greater overhead reduces it. The continuous approximation $k \simeq \sqrt{\Xi/f}$, with a lower bound of one, suppresses these integer thresholds.

Resolution supply and learning speed have different effects. Holding recurrence and the time to routinization fixed, more usable resolutions raise $DR(\pi)$ and can support additional layers. Faster routinization reduces $D = \lambda_{\text{rec}} T_{\text{learn}}$, shortening the time those same resolutions spend at intermediate requirements. It can therefore reduce depth. Synthetic supervision affects depth through the usable ordinary resolutions it generates.

8.4 Joint choice of supervision and depth

Maintain the uniform-path technology on a common nonempty compact implementable supervision interval $\Pi \subseteq [0, \bar{\pi}]$, with continuous R and finite lower semicontinuous C_{common} . For comparative statics, additionally take $D > 0$ and differentiable R with $R'(\pi) > 0$, holding learning speed and all other primitives fixed. A unique optimum at each depth gives an ordering of supervision choices. The stated compactness and lower-semicontinuity conditions ensure attainment.

A single optimally located bridge is at the midpoint. Before overhead, it saves

$$\underbrace{\frac{\gamma \Delta_a M_0}{4}}_{\text{initial resource-matching gain}} + \underbrace{\frac{\gamma \Delta_a DR(\pi)}{4}}_{\text{gain during routinization}}. \quad (49)$$

The learning term values the workload served during partial routinization. The joint problem is

$$\min_{\pi \in \Pi} \left\{ C_{\text{common}}(\pi) + \min_{k \geq 1} \left[\frac{\Xi(\pi)}{k} + f(k-1) \right] \right\}. \quad (50)$$

Corollary 1 (Complementarity between supervision and layers). *The joint minimum in (50) is attained. In the comparative-static regime,*

$$\frac{\partial}{\partial \pi} \left\{ \frac{\Xi(\pi)}{k(k+1)} - f \right\} = \frac{\gamma \Delta_a DR'(\pi)}{2k(k+1)} > 0. \quad (51)$$

Selected depth is weakly increasing in supervision, and the unique supervision optimum is weakly increasing in depth.

Proof of Corollary 1. Differentiating (48) gives (51). For $k' > k$, the cost difference

$$C_{k'}(\pi) - C_k(\pi) = \Xi(\pi)(1/k' - 1/k) + f(k' - k)$$

decreases with π . Cost therefore has decreasing differences in depth and supervision. Comparing revealed-preference inequalities at the two unique supervision optima gives the claimed ordering. Joint minimization of (46) produces (50).

For existence, the bound $0 \leq R(\pi) \leq N_{\text{fam}}$ implies a uniform bound on $\Xi(\pi)$. Every count satisfying $f(k-1) > \sup_{\Pi} \Xi$ is dominated by $k=1$ at every π . It is therefore enough to consider finitely many depths. Each has a lower semicontinuous objective on a compact domain, so a joint minimum is attained. $\square$

The condition for the next interval can be written as the workload threshold

$$M_0 + DR(\pi) > \frac{2fk(k+1)}{\gamma\Delta_a}. \quad (52)$$

or, for $D > 0$, in terms of the resolution threshold

$$\bar{R}_k = \frac{2fk(k+1)}{\gamma\Delta_a D} - \frac{M_0}{D}. \quad (53)$$

Strict savings require $R(\pi) > \bar{R}_k$. If R is continuous and strictly increasing, there is a unique exposure threshold whenever $\bar{R}_k$ lies in its range. Keep k at equality and select $k+1$ above it until the next threshold. A threshold below $\min_{\pi \in \Pi} R(\pi)$ is always exceeded; one above $\max_{\pi \in \Pi} R(\pi)$ is never exceeded. Equality remains a tie.

The joint supervision choice still depends on C_{common} . Resolving a family reduces ideal cycle cost by $\gamma\Delta_a D/2$ but increases residual overprovisioning by $\gamma\Delta_a D/(2k)$. Its net gain before resolution cost is therefore $\gamma\Delta_a D(1 - 1/k)/2$, of which deeper architectures capture more. Improvements in ordinary quality and routine savings after the cycle can benefit every architecture.

This complementarity holds learning speed fixed. Faster routinization shortens T_{learn} and reduces intermediate workload $DR = \lambda_{\text{rec}} T_{\text{learn}} R$, which can lower optimal depth. Synthetic work can also displace ordinary capacity enough to produce $R'(\pi) \leq 0$.

8.5 Locating incentive-bearing supervision

Checks that carry incentives belong where verification inputs are privately chosen. Controlled automated execution still requires evaluation and resources, but not an effort reward. With ordinary tier workload m_{ord} , required synthetic share π , and unit control cost c_A , the synthetic resource bill is

$$c_A \frac{\pi}{1 - \pi} m_{\text{ord}}.$$

Choosing a smaller tier lowers this bill only if the needed exposure, unit cost, and support for ordinary production incentives are unchanged. The reduced-review policy must preserve assessment gaps and the continuation rents needed for verification. Production can remain strategic even after review is automated. Ordinary and synthetic inputs must both be counted, whether rewarded directly or through access.

9 The Gold–Bronze quality–cost frontier

Gold-like quality at Bronze-like cost requires increasingly accurate work as well as declining resource use. The joint limit depends on amortizing preparation while eliminating recurrent overhead in both the validation and worker ledgers. Replication and fresh monitoring identify separate barriers. Throughout this section, fix the workload class and opportunity-cost prices, and compare undiscounted averages over expanding source workloads. Each horizon has its own ex ante plan.

9.1 Successively tighter targets and amortized preparation

For a plan covering N fixed source records, let

$$\bar{L}_N = \frac{1}{N} \sum_{t=1}^N \mathbb{E}[d_\lambda(Y_t, \mathcal{G}_{r_t}(\theta_t))]. \quad (54)$$

This is expected loss relative to the source and the representation ontology. It is not a claim of error-free realized output or of scientific truth in the source. Required omissions remain part of both the loss account and the fixed workload. If every non-Gold release incurs loss of at least $d_{\min} > 0$, then

$$\frac{1}{N} \sum_{t=1}^N \mathbb{P}\{Y_t \notin \mathcal{G}_{r_t}(\theta_t)\} \leq \frac{\bar{L}_N}{d_{\min}}.$$

Absent that separation, vanishing weighted loss need not make non-Gold records vanishingly rare. Errors assigned zero weight receive no certification.

The joint limit is obtained by tightening the tolerance while amortizing the preparation needed at each step.

Let $K_v(\varepsilon), K_m(\varepsilon)$ be nonnegative preparation costs and $h_v(\varepsilon), h_m(\varepsilon)$ nonnegative bounds on recurrent overhead. Preparation includes verification, training, search, and configuration; recurrent costs include audits, consensus, duplicate execution, and maintenance. All inputs are allocated once to the appropriate ledger.

Assumption 3 (Implementable and consistent routinization). For every $\varepsilon \in (0, 1]$, an implementable plan satisfies, for all sufficiently large N ,

$$\begin{aligned} \bar{L}_N &\leq \varepsilon, \\ \mathcal{C}_N^v &\leq K_v(\varepsilon) + N h_v(\varepsilon), \\ \mathcal{C}_N^m &\leq N c_{\text{Br}} + K_m(\varepsilon) + N h_m(\varepsilon), \end{aligned} \quad (55)$$

with $K_v(\varepsilon) + K_m(\varepsilon) < \infty$. Every compared plan satisfies $\mathcal{C}_N^m \geq N c_{\text{Br}}$.

At every tolerance, the plan must jointly satisfy both roles' incentives, participation, funding, and payment bounds. Its compensation and continuation values must be generated by the same policy.

Consistency is an additional condition on the implemented learning technology. For each tolerance, the plan must attain the stated loss bound using the information allowed by Definition 1. Gold labels and a terminal oracle remain unavailable. Learning can change the routine method within its stipulated resource budget.

9.2 Separating validation savings from total-cost convergence

Define the worker excess $\Delta_N^m = \bar{c}_N^m - c_{\text{Br}} \geq 0$. For the *same* quality-constrained plan,

$$\bar{c}_N^{\text{tot}} - c_{\text{Br}} = \bar{c}_N^v + \Delta_N^m. \quad (56)$$

Total cost therefore reaches Bronze if and only if validation cost and worker excess both disappear. Let $\mathcal{C}_N^*(\varepsilon)$ denote the infimum of total unit cost among implementable plans satisfying $\bar{L}_N \leq \varepsilon$.

Theorem 2 (The Gold–Bronze limit in the two resource ledgers). *Under Assumption 3, one can choose implementable plans and tolerances $\varepsilon_N \downarrow 0$, satisfying $\bar{L}_N \leq \varepsilon_N$, with the following properties:*

- (i) *If $h_v(\varepsilon) \rightarrow 0$, then $\bar{c}_N^v \rightarrow 0$.*
- (ii) *If, in addition, $h_m(\varepsilon) \rightarrow 0$, then*

$$\bar{L}_N \rightarrow 0, \quad \bar{c}_N^v \rightarrow 0, \quad \bar{c}_N^{\text{tot}} \rightarrow c_{\text{Br}}, \quad (57)$$

and $\mathcal{C}_N^(\varepsilon_N) \rightarrow c_{\text{Br}}$.*

Proof of Theorem 2. Choose a diagonal sequence that tightens quality only after the horizon is large enough to amortize the chosen plan’s preparation. Set $\varepsilon_j = 1/j$ and $K_j = K_v(1/j) + K_m(1/j)$. For each plan j , choose strictly increasing integers $N_j \rightarrow \infty$ beyond its minimum admissible horizon and with $K_j/N_j \leq 1/j$. Finiteness of every K_j makes these choices possible.

On horizons $N_j \leq N < N_{j+1}$, use the $1/j$ -plan and set $\varepsilon_N = 1/j$. Its expected loss does not exceed $1/j$, and

$$\bar{c}_N^v \leq \frac{1}{j} + h_v(1/j), \quad \bar{c}_N^{\text{tot}} \leq c_{\text{Br}} + \frac{1}{j} + h_v(1/j) + h_m(1/j).$$

The validation inequality proves (i). If both overhead functions vanish, the total-cost inequality converges to the Bronze lower bound and proves (ii). The optimal cost infimum lies between Bronze and the cost of this feasible plan, so it has the same limit. Finally, (13) gives (56). Both terms on its right are nonnegative, so their sum tends to zero exactly when each term does. $\square$

The construction tightens each target after sufficient reuse has amortized its preparation. Both ledgers are evaluated for the same quality-constrained plan.

Appendix C.1 gives a sufficient construction using a reusable majority table built from independent, informative, incentive-compatible votes, without observing Gold. Enforceable Bronze reuse and no recurrent overhead yield $\bar{L}_N \leq 1/N$ with cost gap $O(\log N/N)$. The family set may grow if complete preparation remains sublinear.

9.3 Replication and the single-execution baseline

Suppose a plan uses an expected $m_N \geq 1$ complete worker extractions per source record, each costing at least c_{Br} . All other resource inputs remain nonnegative. This local restriction gives a floor independently of whether preparation can be amortized.

Corollary 2 (Replication keeps the two frontiers apart).

$$\bar{c}_N^{\text{tot}} \geq m_N c_{\text{Br}} + \bar{c}_N^{\text{v}}. \quad (58)$$

If $\liminf_N m_N > 1$, total unit cost cannot converge to c_{Br} , even if both validation cost and fidelity loss tend to zero.

Proof. The worker ledger includes all Nm_N expected executions, irrespective of which workers are paid. Nonnegative costs therefore imply $\mathcal{C}_N^{\text{m}} \geq Nm_N c_{\text{Br}}$. Adding the validation ledger and dividing by N proves the inequality. Taking the lower limit proves the final assertion. $\square$

This lower bound concerns full executions, not several identities submitting a single copied output. Three extractions create a floor of $3c_{\text{Br}}$ even when only one worker wins a prize. Occasional competitions over reusable methods can still satisfy the limit when the combined winning and losing work, training, search, and monitoring is sublinear in the routine volume served. Permanent replication instead calls for a different baseline.

9.4 Fresh monitoring and other obstructions

Persistent demand for fresh monitoring. Suppose a deviation saves $g_{\text{mon}} > 0$, and only an independent audit with probability p_{audit} distinguishes it. If the range of reward exposure is $\bar{B}$, implementation requires $p_{\text{audit}} \bar{B} \geq g_{\text{mon}}$. With audit costs $k_{\text{v}}, k_{\text{m}}$, the corresponding floors are $k_{\text{v}} g_{\text{mon}} / \bar{B}$ and $(k_{\text{v}} + k_{\text{m}}) g_{\text{mon}} / \bar{B}$. Proposition C.2 gives the local assumptions and a sharper bound accounting for informativeness. Separate fresh audits needed on a positive fraction of decisions prevent the total-cost Bronze limit. Outsourcing can eliminate only the curator’s portion of that cost.

By contrast, if the reviewed fraction $r_N \rightarrow 0$ while the synthetic share within that queue stays at $\pi \in (0, 1)$, synthetic volume per source record is $r_N \pi / (1 - \pi)$. At bounded unit costs it vanishes. Routine production outside the queue still requires effective incentives. A positive per-record cost of automated assessment is recurrent overhead unless already included in the maintained baseline.

Continuation rents, novelty, and residual error. Eligibility must retain sufficient value at every history where care is required. Reduced production rewards, delayed assignments, or cheap re-entry can remove that value. Review inputs enter resource costs under either direct or indirect compensation. Novelty at positive density, continuing ordinary audits, intermediate processing, and infrastructure can all leave recurrent overhead. Irreducible error separately rules out zero loss. Unavoidable infrastructure may be included in an explicitly larger baseline. Even when these floors persist, fixed-target resource savings remain valuable.

10 Architecture effects and limits of identification

The preceding design results hold learning opportunities fixed across architectures. Allowing architecture or execution to change learning adds a future cost effect to current resource matching. A separate issue is what observed histories reveal: even an effective incentive policy need not identify latent fidelity. These extensions locate the technological and informational limits of the mechanism.

10.1 When architecture changes learning

If layers change absorption or transfer, index the transition and occupation measure by architecture: $P_{S,\pi}$ and $\mu_{S,\pi}$. As long as the resource used for execution does not itself alter the path, write

$$\mathcal{C}_{\text{exec}}(S, \pi) = \int c(a_S(z)) d\mu_{S,\pi}(z).$$

For an insertion $S' = S \cup \{a_{\text{new}}\}$, with $a_- < a_{\text{new}} < a_+$ between adjacent capacities,

$$\begin{aligned} \mathcal{C}_{\text{exec}}(S, \pi) - \mathcal{C}_{\text{exec}}(S', \pi) &= [c(a_+) - c(a_{\text{new}})]\mu_{S,\pi}((a_-, a_{\text{new}}]) \\ &\quad + \int c(a_{S'}(z)) d[\mu_{S,\pi} - \mu_{S',\pi}](z). \end{aligned} \quad (59)$$

The first component is improved matching along the old path; the second is the effect of a different learning path. The latter is nonnegative under an equal-mass downward first-order stochastic shift. Include changes in discovery and training costs before comparing the total gain with overhead. Layer complementarities can invalidate diminishing returns, the exact-depth rule, and greedy stopping. Appendix D also considers endogenous resolution supply and routing when resource requirements are not directly observed.

When execution itself affects learning, use a sufficient Markov state $\mathbf{z}$. Enlarge that state with histories and promised utilities wherever continuation incentives require them, and let $z(\mathbf{z})$ denote its resource requirement. For a nonempty set of recursively implementable actions $\mathcal{A}_{S,\pi}^{\text{IC}}(\mathbf{z}) \subseteq \{a \in S : a \geq z(\mathbf{z})\}$ and complete action cost $c_{\text{act}}(\mathbf{z}, a; S, \pi)$, the recursion is

$$C_{S,\pi}^{\text{dyn}}(\mathbf{z}) = \min_{a \in \mathcal{A}_{S,\pi}^{\text{IC}}(\mathbf{z})} \left\{ c_{\text{act}}(\mathbf{z}, a; S, \pi) + \delta \int C_{S,\pi}^{\text{dyn}}(\mathbf{z}') P(d\mathbf{z}' \mid \mathbf{z}, a, S, \pi) \right\}. \quad (60)$$

If all technically sufficient actions are implementable and other costs do not depend on the action, the current comparison reduces to $c(a)$. A finite horizon needs terminal values. Bounded costs and $\delta < 1$ give bounded infinite-horizon values. Unlike the common-frontier case, the cheapest current execution may sacrifice valuable later learning.

10.2 Observable effects and latent fidelity

Let O collect reports, synthetic grades, disagreements, resolutions, releases, and observed policy randomization. Two economies may generate identical complete finite-history laws under every maintained policy while having different latent losses. No estimator based on a history of N_{hist} observations, $O_{N_{\text{hist}}}$, can then be consistent in both economies. For a target

gap $\Delta > 0$, its larger expected absolute error is at least $\Delta/2$. Lemma A.5 extends the bound to approximately equivalent history laws.

Policy effects on observables can thus be identified even when their effects on latent quality cannot. Identifying the latter requires justified validity, transport, independent evidence, or subsequent resolution. A human reference adds information only subject to bounds on its incentives and bias.

11 Application: extracting scientific claims and evidence

Scientific extraction turns source passages into structured claims and evidence. Its quality depends on both excluding unsupported content and recovering required content. The application links these two tasks to verification incentives, intermediate resources, and the reuse of difficult judgments.

11.1 What must be verified

A record may contain an estimate, its direction, the population studied, qualifiers, and supporting passages. Fidelity requires these fields to represent the source, not to establish the scientific truth of its claims. A false inclusion can be a value reported where the source supplies none; an omission can be a required population or qualifier absent from the record. The loss in (1) evaluates both.

Evidence from extraction studies motivates this distinction. Gougherty and Clipp (2024) find high accuracy for several discrete and categorical fields but much lower accuracy for some quantitative fields. Konet et al. (2024) show that extraction also depends on the text supplied to the model, with many errors attributable to PDF parsing. These studies motivate field-specific checks and attention to source access.

Synthetic controls can alter an estimate, direction, identifier, or supporting span. Their validity and usefulness are separate. A directional reversal may test careful comparison with the source, but does not reward searching for a missing claim. Omission searches and supported additions therefore need their own payoff consequences, as in Example 1. Reviewing only disagreements also requires a justification for leaving confidently accepted records unchecked.

11.2 Resources and reusable resolutions

Routine extraction, stronger machine assessment, and high-level source inspection provide an interpretation of the resource intensities. An intermediate layer handles cases that no longer need the most expensive resources but cannot yet be processed routinely. Under the common-learning benchmark, it exploits progress rather than causing that progress. Training or distillation costs are charged separately. A layer that changes what can be learned instead falls under the architecture-dependent comparison in Section 10.1.

An ordinary resolution can supply a rule for linking evidence to a claim, preserving a qualifier, or recovering required content. Its resource value is the later workload to which

that rule safely applies. More such resolutions can support additional intermediate layers, whereas faster transfer to routine methods can shorten their useful operating range. The optimal number of layers and their resource intensities therefore depend on the workload and learning path.

11.3 Assignments and evaluation

Assign extraction and current review to different workers, commit submissions before feedback, and assess producers on their submitted work rather than repairs made by reviewers. Concealed controls must preserve ordinary formats, source access, and response obligations. Reward selective verification directly or through future assignments whose net value satisfies (31).

Learning uses ordinary resolutions, with synthetic labels kept attached to planted defects. Evaluation must distinguish changes in synthetic scores, ordinary inclusion and omission losses, and resources required at a maintained quality target. Held-out ordinary cases can inform this evaluation subject to the identification conditions in Section 10.2. Resource comparisons use the same external source workload and count all production and review, including unsuccessful work. This evaluates the joint quality–cost frontier using total resources per source record.

12 Conclusion

The economic value of verification lies partly in what happens after a case is resolved. A judgment produced to assess current work can become an input into later assessment or execution. Verification is then an investment in reusable knowledge. Synthetic challenges can help produce that knowledge when ordinary answers are unknown, provided that they induce behavior that improves ordinary resolutions.

The analysis sharpens two conditions for this mechanism to generate large savings. First, reuse must be sufficient to recover the cost of producing and incorporating a resolution. More usable resolutions can therefore support additional resource layers, while faster routinization can eliminate the need for them: what matters for organizational design is how much work remains at each level of required expertise. Second, reaching Bronze cost requires both validation overhead and worker resources above one routine execution to vanish per output. Nonvanishing replication or monitoring overhead therefore prevents the Gold–Bronze limit even when expert review becomes negligible.

Consider coding records against a fixed medical classification or reviewing documents against an established compliance rule. Resolving an ambiguous classification can inform both later coding and the assessment of coders; resolving a recurring document inconsistency can improve subsequent screening. The question in each case is how to elicit useful judgment and reuse it across decisions that share its factual and procedural basis. The savings come from making that judgment reusable.

A Reference incentives, synthetic validity, and identification

The supporting results distinguish producing a reference, justifying negative labels, rewarding the relevant effort, and identifying ordinary quality.

A.1 A benchmark for producing the reference

Retain the binary-panel setting of Section 4, with its effort costs, agreement payment, and conditional signal structure.

Lemma A.1 (Incentives and error in reference production). *(i) When transfers do not depend on effort or reports, low effort is strictly dominant.*
(ii) Agreement compensation permits the low-effort equilibrium in which all workers submit the focal report z_{foc} . The equilibrium is strict after identifying strategies that differ only at reporting histories never reached.
(iii) If workers exert high effort and report truthfully, majority error obeys

$$\varepsilon_n = \kappa + (1 - \kappa)B_n(p) \leq \kappa + (1 - \kappa)e^{-2n(p-1/2)^2}. \quad (\text{A.1})$$

As $\varepsilon \downarrow \kappa$, the minimum panel resource cost needed for actual error no greater than ε diverges.

Proof. For (i), the transfer is identical under either effort choice. Holding the report fixed, low effort saves c_{eff} , and hence strictly dominates high effort. For (ii), suppose all comparison reports equal z_{foc} . Reporting z_{foc} at low effort pays $T + b$. A different report loses the bonus; high effort incurs c_{eff} without making compensation exceed $T + b$. Every behaviorally distinct deviation is therefore strictly worse. Full contingent strategies can still differ only off path, so strictness is claimed for the reduced game.

For (iii), the common event causes a wrong majority with probability one. Conditional on its complement, majority error is $B_n(p)$, giving the equality in (A.1). Hoeffding's inequality yields the upper bound on the tail, and rearranging it gives (A.2). For every finite odd n , one has $B_n(p) > 0$, although $B_n(p) \rightarrow 0$. Approaching actual error κ therefore requires $n \rightarrow \infty$, which makes the panel cost $n\chi$ diverge. $\square$

To attain a desired error bound $\varepsilon \in (\kappa, 1)$, it is sufficient to choose the next admissible odd panel size at or above

$$n \geq \frac{\log[(1 - \kappa)/(\varepsilon - \kappa)]}{2(p - 1/2)^2}. \quad (\text{A.2})$$

A.2 Sparse admissibility and contaminated controls

Use the predicate-specific response space and mutation kernel introduced in Section 5. Two comparisons separate a mutation's validity from its effect on incentives.

Sparse admissibility. Maintain a finite response space $\mathcal{Y}_r$ and, for every maintained state θ ,

$$|\mathcal{G}_r^{\text{test}}(\theta)| \leq m_r, \quad \sup_{\hat{y}, y \in \mathcal{Y}_r} M_r(y \mid \hat{y}) \leq \bar{p}_r.$$

Neither bound requires observing an admissible value.

Lemma A.2 (A mutation bound from sparse admissibility). *The kernel M_r is ϵ_r -valid for $\epsilon_r = \min\{1, m_r \bar{p}_r\}$.*

Proof. For any seed record $\hat{y}$,

$$M_r(\mathcal{G}_r^{\text{test}}(\theta) \mid \hat{y}) = \sum_{y \in \mathcal{G}_r^{\text{test}}(\theta)} M_r(y \mid \hat{y}) \leq |\mathcal{G}_r^{\text{test}}(\theta)| \bar{p}_r \leq m_r \bar{p}_r.$$

Since a probability is also bounded above by one, complementing this bound gives the claimed lower bound on negative-label validity. $\square$

Contaminated controls. Fix policies H, L . An exogenous validity event has probability $1 - \epsilon_{\text{act}}$, with $0 \leq \epsilon_{\text{act}} \leq \epsilon \leq 1$. The conditional pass-probability gap on valid controls is at least $\Delta p_{\text{val}} \in [0, 1]$. Leave behavior on invalid controls unrestricted unless the stated subcase restricts its gap.

Lemma A.3 (How contamination changes the incentive gap).

$$p_A^H - p_A^L \geq (1 - \epsilon) \Delta p_{\text{val}} - \epsilon. \quad (\text{A.3})$$

The bound is sharp. A nonnegative invalid-control gap gives the stronger bound $(1 - \epsilon) \Delta p_{\text{val}}$.

Proof. Decompose the pass difference by validity. On the invalid component, the difference between two probabilities is at least -1 , so the total is bounded below by $(1 - \epsilon_{\text{act}}) \Delta p_{\text{val}} - \epsilon_{\text{act}}$. This expression decreases in ϵ_{act} . It is attained by valid-control pass probabilities $\Delta p_{\text{val}}, 0$, invalid-control probabilities $0, 1$, and $\epsilon_{\text{act}} = \epsilon$. Requiring a nonnegative gap on the invalid component gives the stronger bound. An effort-independent invalid component contributes zero. $\square$

A.3 Score gradients and implementable effort

Fix a contract and an interior target in a convex effort domain. Hold ξ and reporting continuation fixed, and assume differentiable, strictly concave utility in (6). Define the score Jacobian and marginal settlement weights by

$$\mathcal{J}_j(e_j) = D_{e_j} \mathbf{s}_j(e_j) \in \mathbb{R}^{K_{\text{score}} \times K_e}, \quad \Omega_j(e_j) = \nabla_{\mathbf{s}} \Gamma_j(\mathbf{s}_j(e_j); \xi) \in \mathbb{R}^{K_{\text{score}}}.$$

For the coefficient-choice comparison, permit unrestricted affine weights in $\mathbb{R}^{K_{\text{score}}}$ and consider solvability of the first-order system.

Lemma A.4 (The behavioral span needed to implement effort). *The target $\hat{e}_j$ is the unique conditional best response exactly when*

$$\mathcal{J}_j(\hat{e}_j)^\top \Omega_j(\hat{e}_j) = \nabla C_j(\hat{e}_j). \quad (\text{A.4})$$

In the coefficient-choice comparison, the first-order equation is solvable if and only if

$$\nabla C_j(\hat{e}_j) \in \text{range}(\mathcal{J}_j(\hat{e}_j)^\top).$$

Proof. Differentiating (6) while fixing ξ and reporting continuation gives

$$\nabla U_j(e_j) = \mathcal{J}_j(e_j)^\top \Omega_j(e_j) - \nabla C_j(e_j).$$

Strict concavity on the convex domain makes the first-order condition both necessary and sufficient for a unique interior conditional optimum. Every score-mediated marginal reward belongs to the range of $\mathcal{J}_j^\top$, establishing necessity of the range restriction. Conversely, if $\nabla C_j(\hat{e}_j) = \mathcal{J}_j(\hat{e}_j)^\top \omega_{\text{pay}}$, the affine rule $\Gamma_j(\mathbf{s}) = \omega_{\text{pay}}^\top \mathbf{s}$ supplies the required gradient. It globally implements the target provided the resulting utility also satisfies the maintained strict concavity. $\square$

Restricting affine weights to be nonnegative replaces the linear range by the nonnegative cone. For an actual admissible coefficient set $\mathcal{W}_j$, the attainable first-order incentives are $\mathcal{J}_j(\hat{e}_j)^\top \mathcal{W}_j$. Payment limits can restrict this to a proper bounded subset of the cone. Participation and global curvature must still be checked separately.

A.4 Identification without an oracle

For a fixed policy, denote by $P_{j, N_{\text{hist}}}$ the distribution of complete observable history $O_{N_{\text{hist}}}$ in economy j , including randomized policy choices. Let L_0, L_1 be scalar loss targets that do not vary with sample size, with $\Delta = |L_1 - L_0| > 0$. Use the convention $\text{TV}(P, Q) = \sup_A |P(A) - Q(A)|$.

Lemma A.5 (An identification bound without an oracle). *Any estimator using only $O_{N_{\text{hist}}}$ satisfies*

$$\max_{j=0,1} \mathbb{E}_j |\hat{L}_{N_{\text{hist}}} - L_j| \geq \frac{\Delta}{2} \{1 - \text{TV}(P_{0, N_{\text{hist}}}, P_{1, N_{\text{hist}}})\}. \quad (\text{A.5})$$

If the complete observable laws agree at every finite N_{hist} for every maintained policy, loss is not point identified. No such estimator can be consistent in both economies, and at every finite sample size its larger expected absolute error is at least $\Delta/2$.

Proof. When the laws agree, an $O_{N_{\text{hist}}}$ -based estimator has the same distribution in both economies. It cannot converge in probability to two distinct constants. For targets L_0, L_1 , the pointwise inequality $|\hat{L} - L_0| + |\hat{L} - L_1| \geq |L_1 - L_0|$ integrated under the common law gives the finite-sample risk bound. For differing laws, let p_0, p_1 be their densities with respect to $P_{0, N_{\text{hist}}} + P_{1, N_{\text{hist}}}$. Integrate the triangle inequality against $\min\{p_0, p_1\}$. That measure has mass $1 - \text{TV}(P_{0, N_{\text{hist}}}, P_{1, N_{\text{hist}}})$, and the sum of the two risks is at least the resulting integral. This proves (A.5). $\square$

B Policy-consistent continuation and separate-role implementation

Eligibility rents must be generated by the same policy whose incentives they support. The continuation equations check this consistency, while a finite-outcome comparison tests whether bounded payments can deter all specified deviations at once.

B.1 Histories, assignment policies, and deviations

Maintain perfect recall, disjoint current roles, and commitment to assignment, assessment, release, synthetic generation and grading, score updates, eligibility, payments, and exit or re-entry. A finite horizon has policy-specified terminal values; on an infinite horizon, assume bounded per-period payoffs and continuation values with $0 < \delta_w < 1$. History h includes roles, scores, assignments, private information, and payoff-relevant beliefs. Model versions, other workers' reports, and pool conditions belong in that history whenever they affect assessment or opportunities.

Fix the other workers' strategies and consistent beliefs. Let $\mathcal{A}_i(h)$ be the worker's feasible strategies for the current assignment, including contingent acquisition, reporting, and nonresponse. If no assignment is offered, the set instead includes available waiting and entry decisions. A deviation returns to the recommendation after the current assignment. Let $P_i(do, dh' \mid \varphi, h)$ give the resulting current-outcome and next-history law, $t_i(o)$ the current transfer, and $C_i(\varphi; h)$ the expected resource cost borne by the worker. Both the law and cost include every within-assignment contingency. Define

$$W_i(\varphi, h; V_i) = \int [t_i(o) + \delta_w V_i(h')] P_i(do, dh' \mid \varphi, h) - C_i(\varphi; h). \quad (\text{B.1})$$

For recommended assignment strategy $\varphi_i^*(h)$, policy consistency requires

$$V_i(h) = W_i(\varphi_i^*(h), h; V_i), \quad V_i(h) \geq W_i(\varphi, h; V_i) \quad \text{for every } \varphi \in \mathcal{A}_i(h), \quad (\text{B.2})$$

and $V_i(h) \geq V_i^{\text{exit}}(h)$ wherever exit is feasible. Re-entry strategies include their actual costs, delays, and treatment of previous scores. The conditions apply at every relevant history, with Bayesian updating where appropriate.

Backward induction establishes sufficiency on a finite horizon. For bounded discounted payoffs, iterating the assignment inequalities over any finite sequence of deviations leaves a discounted terminal-value difference. That difference vanishes as the number of periods grows, so no multi-period unilateral deviation improves utility. Imposing the conditions on all workers with consistent beliefs and assessment laws supports sequentially optimal behavior under the committed policy.

B.2 The net rents supplied by eligibility

An eligibility threshold maps updated history into access status. The summary values V_E, V_I are sufficient only if they capture every strategy-dependent continuation effect. Otherwise, retain the full next-history law in (B.1). Ordinary reports may change other workers' production or future pool conditions without changing the reporter's eligibility.

Access value is net of production effort, waiting, and later verification, and includes the possibility of recovery from ineligibility. These values must be generated by a funded policy and recalculated when prizes change. Future transfers enter the resource-financing account when paid, rather than being counted again through their incentive value.

Cheap re-entry can nullify exclusion. Score updates without payoff consequences can leave current care unrewarded. A terminal costly review needs current compensation or another actual consequence.

B.3 The observable distinctions required by incentives

For either role, collect contractible outcomes and continuation states in o . With finitely many outcomes and fixed values, put $b_o = t_i(o) + \delta_w V_i(h'(o))$. The target φ^* requires

$$\sum_o [P_i(o \mid \varphi^*, h) - P_i(o \mid \varphi, h)] b_o \geq C_i(\varphi^*; h) - C_i(\varphi; h) \quad \text{for every } \varphi. \quad (\text{B.3})$$

Current payments must satisfy their bounds, and continuation values must be policy-consistent. A lower-cost strategy with the same complete observable law cannot be deterred. Equality of marginal means alone, however, need not imply equality of that law. All relevant deviations must be included.

Fix a history, finite current-outcome set $\mathcal{O}$, target φ^* , and complete deviations $\varphi^1, \dots, \varphi^J$, with $J \geq 1$. Let their probability vectors be p^*, p^j , and let D_j be the target's fixed, policy-consistent continuation advantage relative to φ^j . Set $g_j = C(\varphi^*) - C(\varphi^j) - D_j$ and impose the finite current-transfer bound $\bar{B} \geq 0$. The following test characterizes the incentive constraints conditional on the continuation policy. Participation and shared-pool feasibility are imposed separately. For finite laws, use $\text{TV}(p, q) = \frac{1}{2} \sum_o |p_o - q_o|$, and let Δ_J denote the probability simplex over the J deviations.

Lemma B.1 (Mixture tests for bounded-pay implementation). *A common transfer vector $\mathbf{t}_{\text{pay}} \in [0, \bar{B}]^{\mathcal{O}}$ satisfies all incentive constraints if and only if, for every $\lambda^{\text{mix}} \in \Delta_J$,*

$$\sum_{j=1}^J \lambda_j^{\text{mix}} g_j \leq \bar{B} \text{TV} \left(p^*, \sum_{j=1}^J \lambda_j^{\text{mix}} p^j \right). \quad (\text{B.4})$$

Proof. For any zero-sum $\mathbf{d}_{\text{pay}}$, $\max_{0 \leq t_o^{\text{pay}} \leq \bar{B}} \mathbf{d}_{\text{pay}}^\top \mathbf{t}_{\text{pay}} = \bar{B} \sum_o [d_o^{\text{pay}}]_+$. The maximum is attained by paying the cap wherever $d_o^{\text{pay}} > 0$ and zero elsewhere. Feasibility of the incentive system is equivalent to $\max_{\mathbf{t}_{\text{pay}}} \min_j \{ (p^* - p^j)^\top \mathbf{t}_{\text{pay}} - g_j \} \geq 0$. Finite-dimensional minimax applies to the compact transfer box and probability simplex, making this maximum equal to

$$\min_{\lambda^{\text{mix}} \in \Delta_J} \left\{ \bar{B} \text{TV} \left(p^*, \sum_j \lambda_j^{\text{mix}} p^j \right) - \sum_j \lambda_j^{\text{mix}} g_j \right\}.$$

The equivalence, including the one-deviation formula, follows. $\square$

With one deviation, the greatest attainable current-payment advantage is $\bar{B} \text{TV}(p^*, p^1)$.

Distinguishing the target from each deviation separately is not sufficient. Take $p^* = (1/2, 1/2)$, $p^1 = (1, 0)$, $p^2 = (0, 1)$, and $g_1 = g_2 = g > 0$. Either deviation by itself can

be deterred if $\bar{B} \geq 2g$. Their equally weighted mixture reproduces the target distribution, however, and violates (B.4). No common transfer therefore deters both. Participation, pool feasibility, and continuation consistency remain additional conditions, all requiring renewed evaluation after a repeated payment policy changes.

C A sufficient frontier construction and a monitoring obstruction

Reusable judgments can meet the joint quality–cost limit in a restricted finite-family setting. Fresh incentive deficits create the opposing case: even fully reusable knowledge cannot remove the cost of repeatedly distinguishing care from shortcuts. The construction and the monitoring bound separate these two requirements.

C.1 Attainment with finitely many reusable families

There are $F \geq 1$ fixed recurring source-task families, each having a fixed latent binary answer. Executing a routine method with the correct family answer has zero loss; all output losses are at most $L_{\max} > 0$. For every odd n , suppose a complete implementable preparation procedure obtains n votes for each family without observing Gold. Conditional on the latent answers, votes within a family are independent and each is correct with probability at least $p > 1/2$. Preparation uses at most $K_0 + c_{\text{vote}}Fn$ resources, including setup, controls, and worker effort. The majority table can then be executed for exactly c_{Br} per record, with no further recurrent cost or fresh hidden-action deficit. This execution is enforceable within the complete plan.

For $\varepsilon \in (0, 1]$, choose n_ε to be the smallest positive odd integer no smaller than

$$\max \left\{ 1, \frac{\log(F L_{\max}/\varepsilon)}{2(p - 1/2)^2} \right\}.$$

Proposition C.1 (A finite-family construction attaining the frontier). *Using n_ε votes per family gives a majority-table plan with average expected loss at most ε at every production horizon. Preparation requires $O(\log(1/\varepsilon))$ resources. Choosing $\varepsilon_N = 1/N$ yields*

$$\bar{L}_N \leq \frac{1}{N}, \quad 0 \leq \bar{c}_N^{\text{tot}} - c_{\text{Br}} \leq \frac{K_0 + c_{\text{vote}}Fn_{1/N}}{N} = O\left(\frac{\log N}{N}\right).$$

Validation cost per record also tends to zero.

Proof. Conditional Hoeffding bounds the probability that a family’s majority is wrong by $\exp[-2n(p - 1/2)^2]$. A union bound puts the probability of any erroneous table entry below $F \exp[-2n(p - 1/2)^2]$. If every entry is correct, all routine releases have zero loss; otherwise each release loses at most $L_{\max}$. The choice of n_ε therefore bounds average expected loss by ε , even when production selects families adaptively. Rounding upward to an odd integer adds fewer than two votes. Combining the preparation bound with zero recurrent overhead gives the cost rates. Nonnegative validation expenditure is bounded by total preparation expenditure, so its unit cost vanishes as well. $\square$

The construction combines informative, incentive-compatible preparation with enforceable routine reuse. Any additional assessment needed to sustain reuse must be charged to recurrent costs.

The argument also allows a horizon-dependent catalogue with F_N families if the informative-vote, complete preparation-cost, and enforceable c_{Br} -execution bounds hold uniformly. Costs of discovering and addressing the expanded catalogue must be included. Suppose the families are available for preparation before production. For $\varepsilon_N = 1/N$, choose the least positive odd n_N at or above

$$\max \left\{ 1, \frac{\log(F_N L_{\max} N)}{2(p - 1/2)^2} \right\}.$$

Then

$$\bar{L}_N \leq N^{-1}, \quad 0 \leq \bar{c}_N^{\text{tot}} - c_{\text{Br}} \leq \frac{K_0 + c_{\text{vote}} F_N n_N}{N}.$$

The condition $F_N \log(N F_N) = o(N)$ consequently suffices for the joint limit. Uncontrolled arrivals, rising lookup costs, and costs incurred before a family is resolved must be added to the preparation or recurrent accounts.

C.2 A lower bound from recurring monitoring

The frontier requires recurrent supervision expenditure per record to vanish. A local incentive comparison can rule this out without treating an audit as an observed correct ordinary answer. Consider a decision whose desired behavior costs $g_{\text{mon}} > 0$ more than an available deviation. An audit is drawn independently of behavior with probability p_{audit} . In the absence of an audit, the two behaviors have identical distributions of all reward-relevant observations and continuation states. The total reward exposure attributable to the decision, including permitted continuation exposure, has range at most $0 < \bar{B} < \infty$. A zero range cannot cover a positive deficit. Keep all other costs and opportunities fixed.

Let $d_A \in [0, 1]$ bound the total-variation distance between the audited laws of the complete reward-relevant outcome. An audit consumes at least $k_v \geq 0$ validation resources and $k_m \geq 0$ worker resources. For a workload-wide bound, suppose these restrictions apply to at least a fraction $\alpha_{\text{mon}} > 0$ of decisions, each requiring its own nonreusable audit.

Proposition C.2 (A persistent monitoring cost floor). *Implementation of the specified local decision requires*

$$p_{\text{audit}} \bar{B} d_A \geq g_{\text{mon}}. \quad (\text{C.1})$$

When $d_A > 0$, expected audit expenditure per decision is at least $k_v g_{\text{mon}} / (\bar{B} d_A)$ for validation and $(k_v + k_m) g_{\text{mon}} / (\bar{B} d_A)$ in total. The workload-wide floors are α_{mon} times these bounds. For $d_A = 0$ or $g_{\text{mon}} > \bar{B} d_A$, even certain auditing cannot implement the desired behavior.

Proof. Without an audit, expected reward is identical for the two behaviors. Conditional on an audit, integrating a payoff of range at most $\bar{B}$ against the two outcome laws bounds the expectation gap by $\bar{B}$ times their total-variation distance. The expected incentive advantage is thus at most $p_{\text{audit}} \bar{B} d_A \leq p_{\text{audit}} \bar{B}$, and it must cover g_{mon} . These are the two incentive bounds.

Expected audit inputs are at least $p_{\text{audit}}k_v$ in the validation ledger and $p_{\text{audit}}(k_v + k_m)$ in the combined ledger. Substitution of the lower bound on p_{audit} gives the cost floors. Summing costs of the separate audits over the stated fraction of decisions gives the density-adjusted bounds. $\square$

When there is no informative bound on audited laws, taking $d_A = 1$ gives $p_{\text{audit}}\bar{B} \geq g_{\text{mon}}$ and the weaker floors $k_v g_{\text{mon}}/\bar{B}$ and $(k_v + k_m)g_{\text{mon}}/\bar{B}$.

The necessity result addresses a local deficit not already covered by ordinary scoring. Shared consequences require analysis of the joint reward exposure and audit technology, rather than addition of local lower bounds. Increasing continuation stakes, decreasing gains from deviation, or assessment built into execution may alter the obstruction. Suppose per-period transfers have a common cap $\bar{t}$, effort costs are nonnegative, permanent exit is free and worth zero, and the common discount factor satisfies $\delta_w < 1$. Participation then implies $0 \leq V_i(h) \leq \bar{t}/(1 - \delta_w)$. Current transfer plus discounted continuation has range no greater than $\bar{t}/(1 - \delta_w)$, furnishing a uniform $\bar{B}$. Uniformity can fail if discount factors tend to one or if payment and exit-value bounds change across plans.

Audit work additional to routine execution is overhead above Bronze. Outsourcing may set $k_v = 0 < k_m$, removing validation expenditure but preserving the total-cost floor. Accounting must identify both the function performed and the party that bears its resource cost.

D Learning paths and resource design

These results establish how usable resolutions propagate, when endogenous resolution supply remains stable, and how resource savings justify layers. The final comparison replaces directly observed requirements with posterior-risk routing.

D.1 Stable propagation along learning paths

Fix the architecture and usable-resolution input of (34), maintaining $\mathbf{A}_S \geq 0$, $\mathbf{b}_S \geq 0$, and $\text{spr}(\mathbf{A}_S) < 1$. Where derivatives are reported, assume R is differentiable. These standard stability restrictions are separate from the effort and ordinary-quality conditions that determine the resolution input.

Lemma D.1 (The multiplier from repeated learning). *The unique steady state of (34) is*

$$x^*(\pi) = (I - \mathbf{A}_S)^{-1} \mathbf{b}_S R(\pi),$$

and its derivative is

$$\frac{dx^*(\pi)}{d\pi} = \sum_{k=0}^{\infty} \mathbf{A}_S^k \mathbf{b}_S R'(\pi). \quad (\text{D.1})$$

If $R'(\pi) > 0$, a component improves strictly if and only if a positive path connects the direct resolution input to that component.

Proof. The spectral-radius condition makes $I - \mathbf{A}_S$ invertible and gives the Neumann series $(I - \mathbf{A}_S)^{-1} = \sum_{k \geq 0} \mathbf{A}_S^k$. Solving the fixed-point equation and differentiating in π yields (D.1). Since the matrices and inputs are nonnegative, a component is strictly positive precisely when it is reachable along a positive matrix path. $\square$

D.2 Endogenous resolution supply

Take a funded, implementable policy family and $\mathbf{A} \geq 0, \mathbf{b} \geq 0$ with $\text{spr}(\mathbf{A}) < 1$. Let capability evolve as $x_{t+1} = \mathbf{A}x_t + \mathbf{b}\mathcal{R}(x_t, \pi)$. Assume $\mathcal{R}(x, \pi) \geq 0$ is continuously differentiable and nondecreasing in π , with the componentwise bound $0 \leq \nabla_x \mathcal{R}(x, \pi) \leq \mathbf{d}_{\text{fb}}$ for a common $\mathbf{d}_{\text{fb}} \geq 0$ throughout $\mathbb{R}_+^{L_{\text{lay}}} \times \Pi$. Define $\boldsymbol{\eta}^{\text{learn}} = (\mathbf{I} - \mathbf{A})^{-1}\mathbf{b}$, and impose the small-gain condition $\mathbf{d}_{\text{fb}}^\top \boldsymbol{\eta}^{\text{learn}} < 1$. It bounds the additional feedback from capability to resolution supply.

Proposition D.1 (Stable learning with endogenous resolutions). *For each π , the stated feedback system has one globally attracting steady state. On a differentiable parameter branch,*

$$\frac{dx^*}{d\pi} = \frac{\boldsymbol{\eta}^{\text{learn}} \partial_\pi \mathcal{R}(x^*, \pi)}{1 - \nabla_x \mathcal{R}(x^*, \pi)^\top \boldsymbol{\eta}^{\text{learn}}} \geq 0. \quad (\text{D.2})$$

The response of component j is strict exactly when both $\eta_j^{\text{learn}} > 0$ and $\partial_\pi \mathcal{R}(x^, \pi) > 0$.*

Proof. Let $\mathbf{M}_{\text{fb}} = (\mathbf{I} - \mathbf{A})^{-1}$ and $\mathbf{H}_{\text{fb}} = \mathbf{A} + \mathbf{b}\mathbf{d}_{\text{fb}}^\top$. For the positive vector

$$\mathbf{w}_{\text{fb}} = \mathbf{M}_{\text{fb}} \mathbf{1} + \boldsymbol{\eta}^{\text{learn}} \frac{\mathbf{d}_{\text{fb}}^\top \mathbf{M}_{\text{fb}} \mathbf{1}}{1 - \mathbf{d}_{\text{fb}}^\top \boldsymbol{\eta}^{\text{learn}}} > 0$$

one has $\mathbf{H}_{\text{fb}} \mathbf{w}_{\text{fb}} = \mathbf{w}_{\text{fb}} - \mathbf{1} < \mathbf{w}_{\text{fb}}$. By the mean-value bound on $\mathcal{R}$, the update is a contraction on the closed nonnegative orthant under $\|x\|_{\mathbf{w}_{\text{fb}}} = \max_j |x_j|/(\mathbf{w}_{\text{fb}})_j$, with modulus at most $\max_j (\mathbf{H}_{\text{fb}} \mathbf{w}_{\text{fb}})_j / (\mathbf{w}_{\text{fb}})_j < 1$. A unique fixed point exists, and every trajectory converges to it. The fixed point obeys $x^* = \boldsymbol{\eta}^{\text{learn}} \mathcal{R}(x^*, \pi)$. Differentiate this equation. The denominator in (D.2) is no smaller than $1 - \mathbf{d}_{\text{fb}}^\top \boldsymbol{\eta}^{\text{learn}} > 0$, proving both the formula and its sign. $\square$

When capability changes review queues, eligibility rents, or reporting behavior, those effects enter $\mathcal{R}$ and its derivative bound.

D.3 A local identity for adding a layer

Maintain Assumptions 1 and 2, fixing supervision and the potential workload μ_π . Let $a_- < a_+$ be adjacent installed capacities and propose $a_{\text{new}} \in (a_-, a_+)$, with full fixed and interface cost f_{new} .

Lemma D.2 (The primitive condition for a profitable insertion). *Installing a_{new} lowers total cost if and only if*

$$[c(a_+) - c(a_{\text{new}})] \mu_\pi((a_-, a_{\text{new}}]) > f_{\text{new}}. \quad (\text{D.3})$$

Equality leaves the two configurations equally costly.

Proof. Before a_{new} is installed, requirements in $(a_-, a_{\text{new}}]$ are served at cost $c(a_+)$. After installing a_{new} , they are served at cost $c(a_{\text{new}})$; every other assignment is unchanged. Integrate this difference using (10) and deduct f_{new} to obtain the result. $\square$

Future decisions whose learning paths reach $(a_-, a_{\text{new}}]$ are already part of the occupation measure, so their savings enter this same calculation.

D.4 Globally optimal partitions

Retain the common learning frontier, finite workload, continuous increasing execution cost, and positive per-layer overhead of Section 8. The proof allows atoms and coincident cutoffs; scaled requirements are used only for the final comparative static.

Proof of Proposition 4. For fixed k , consider the compact ordered simplex of cutoffs, including coincident ones. Along any converging sequence of configurations, the limiting cost of the least feasible installed intensity at requirement z is no larger than the lower limit of the corresponding sequence of costs. A cutoff that reaches z permits a downward cost jump, not an upward jump. Continuity of c , nonnegativity, and Fatou’s lemma make integrated overprovisioning lower semicontinuous. Hence a minimum exists. Coincident cutoffs that represent redundant layers can be removed to save their fixed costs. Since interval costs add, conditioning on the penultimate cutoff gives (40). Removing the common ideal-resource cost from the objective gives (41). Because $E_k \geq 0$, any count with $f(k-1) > E_1$ is dominated by the endpoint-only choice $k = 1$. Comparing optimality inequalities at two values of f shows that the selected interval count k cannot rise when layer overhead rises.

For $a_- \leq a_+ \leq a_{\text{mid}} \leq a_{\text{end}}$, subtraction gives the Monge relation

$$H_\pi(a_-, a_{\text{end}}) + H_\pi(a_+, a_{\text{mid}}) - H_\pi(a_-, a_{\text{mid}}) - H_\pi(a_+, a_{\text{end}}) = [c(a_{\text{end}}) - c(a_{\text{mid}})]\mu_\pi((a_-, a_+)) \geq 0. \quad (\text{D.4})$$

Its partition implication can be established without approximating by a grid. Take optimal partitions $p_0, \dots, p_{k-1}$ and $q_0, \dots, q_{k+1}$ with identical endpoints. Choose the first $i \in \{1, \dots, k-1\}$ satisfying $p_i \geq q_{i+1}$. Such an index exists because $p_{k-1} = \bar{a}$, and it obeys $p_{i-1} \leq q_i \leq q_{i+1} \leq p_i$. Replace $(p_{i-1}, p_i]$ and $(q_i, q_{i+1}]$ by $(p_{i-1}, q_{i+1}]$ and $(q_i, p_i]$, exchanging the subsequent tails. This produces

$$(p_0, \dots, p_{i-1}, q_{i+1}, \dots, q_{k+1}), \quad (q_0, \dots, q_i, p_i, \dots, p_{k-1}).$$

The new interval counts are $(i-1) + 1 + (k-i) = k$ and $i + 1 + (k-1-i) = k$. Every unaffected interval has its original cost. By (D.4), total cost for the new pair is no greater than $E_{k-1} + E_{k+1}$, while optimality at count k bounds it below by $2E_k$. This proves discrete convexity, including the cases of atoms and coincident cutoffs. It is the usual Monge path exchange; Wan (2024) gives a discrete shortest-path formulation. Hence $f - (E_k - E_{k+1})$ is nondecreasing in k . The nonnegative sequence E_k is nonincreasing, so its successive improvements tend to zero. The first nonnegative total-cost increment therefore exists and yields (42).

After normalizing requirements to $\tilde{z}$, every execution-cost difference equals $\gamma\Delta_a$ times its normalized counterpart. Additional cutoffs cannot worsen normalized overprovisioning, giving $\mathcal{E}_{k+1}^{\text{norm}} \leq \mathcal{E}_k^{\text{norm}}$. In $\gamma\Delta_a\mathcal{E}_k^{\text{norm}} + f(k-1)$, an increase in $\gamma\Delta_a/f$ thus weakly favors larger k . The same revealed-preference comparison, together with the specified tie-break, proves the scaled-heterogeneity result. $\square$

D.5 Posterior-risk routing without observed requirements

Posterior-risk comparisons can replace the observed-requirement rule (4). When requirements are too noisy to observe directly, evaluate complete actions with (14). Specialize the observable

task type r to scalar posterior risk. For a fixed active set with routine production P, an intermediate technology B, and high-level review H, define the adjacent cost gaps

$$G_{PB}(r) = J_P(r) - J_B(r), \quad G_{BH}(r) = J_B(r) - J_H(r).$$

Let r vary over a scalar interval. Assume G_{PB} and G_{BH} are continuous and strictly increasing, with interior roots r_{PB} and r_{BH} . Their ordering determines whether the bridge serves a nonempty risk region.

Proposition D.2 (Threshold routing with three technologies). *If $r_{PB} < r_{BH}$, pointwise minimum-cost routing is*

$$\psi^*(r) = \begin{cases} P, & r < r_{PB}, \\ B, & r_{PB} < r < r_{BH}, \\ H, & r > r_{BH}. \end{cases}$$

Adjacent technologies tie at their respective thresholds. If $r_{PB} > r_{BH}$, the bridge is strictly dominated at every risk value. The lower cost envelope then moves directly from P to H at a unique $r_{PH} \in (r_{BH}, r_{PB})$. If the roots coincide, all three tie at that cutoff, with P cheaper below it and H cheaper above it.

Proof. Below both roots, both adjacent differences are negative, so $J_P < J_B < J_H$. Between correctly ordered roots, the bridge costs less than either endpoint. Above both roots, high-level adjudication is cheapest. When the roots are reversed, some endpoint is cheaper than the bridge in every region. The endpoint gap $J_P - J_H = G_{PB} + G_{BH}$ is strictly increasing. For reversed roots it is negative at r_{BH} and positive at r_{PB} , so continuity gives exactly one crossing between them. For coincident roots, it vanishes at their common threshold. $\square$

A nonempty bridge region carries workload only if the risk distribution places positive mass there. Installation is worthwhile only if savings over that workload cover fixed and interface costs.

References

- Arrow, K. J. (1962). The economic implications of learning by doing. *Review of Economic Studies*, 29(3), 155–173.
- Barr, E. T., Harman, M., McMinn, P., Shahbaz, M., and Yoo, S. (2015). The oracle problem in software testing: A survey. *IEEE Transactions on Software Engineering*, 41(5), 507–525.
- Caliendo, L. and Rossi-Hansberg, E. (2012). The impact of trade on organization and productivity. *Quarterly Journal of Economics*, 127(3), 1393–1467.
- Checchio, A., Bates, J., and Demartini, G. (2020). Adversarial attacks on crowdsourcing quality control. *Journal of Artificial Intelligence Research*, 67, 375–408.
- Dasgupta, A. and Ghosh, A. (2013). Crowdsourced judgement elicitation with endogenous proficiency. In *Proceedings of the 22nd International Conference on World Wide Web*, 319–330.

- Fang, K., Powell, G. E., Painter, J. L., Bate, A., and Li, M. L. (2026). The right amount of wrong: Error injection for AI oversight. *SSRN preprint*.
- Gao, X. A., Wright, J. R., and Leyton-Brown, K. (2019). Incentivizing evaluation with peer prediction and limited access to ground truth. *Artificial Intelligence*, 275, 618–638.
- Garicano, L. (2000). Hierarchies and the organization of knowledge in production. *Journal of Political Economy*, 108(5), 874–904.
- Garicano, L. and Rossi-Hansberg, E. (2012). Organizing growth. *Journal of Economic Theory*, 147(2), 623–656.
- Gougherty, A. V. and Clipp, H. L. (2024). Testing the reliability of an AI-based large language model to extract ecological information from the scientific literature. *npj Biodiversity*, 3, Article 13.
- Holmström, B. (1979). Moral hazard and observability. *Bell Journal of Economics*, 10(1), 74–91.
- Holmström, B. and Milgrom, P. (1991). Multitask principal–agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics, and Organization*, 7(Special Issue), 24–52.
- Konet, A., Thomas, I., Gartlehner, G., Kahwati, L., Hilscher, R., Kugley, S., Crotty, K., Viswanathan, M., and Chew, R. (2024). Performance of two large language models for data extraction in evidence synthesis. *Research Synthesis Methods*, 15(5), 818–824.
- Lazear, E. P. and Rosen, S. (1981). Rank-order tournaments as optimum labor contracts. *Journal of Political Economy*, 89(5), 841–864.
- Miller, N., Resnick, P., and Zeckhauser, R. (2005). Eliciting informative feedback: The peer-prediction method. *Management Science*, 51(9), 1359–1373.
- Oleson, D., Sorokin, A., Laughlin, G. P., Hester, V., Le, J., and Biewald, L. (2011). Programmatic Gold: Targeted and scalable quality assurance in crowdsourcing. In *Human Computation: Papers from the 2011 AAAI Workshop*, 43–48.
- Shapiro, C. and Stiglitz, J. E. (1984). Equilibrium unemployment as a worker discipline device. *American Economic Review*, 74(3), 433–444. <https://www.jstor.org/stable/1804018>.
- Wan, J. Z. (2024). Finding a shortest M -link path in a Monge directed acyclic graph. *arXiv preprint*, arXiv:2408.00227.
- Witkowski, J., Freeman, R., Vaughan, J. W., Pennock, D. M., and Krause, A. (2023). Incentive-compatible forecasting competitions. *Management Science*, 69(3), 1354–1374.