StructPay: Joint Workflow and Reward Design for Reliable Agent Services

Anonymous authors
Paper under double-blind review

Abstract

Workflow dependence can make reliable effort expensive, while heterogeneous services can make a common reward inefficient. We introduce StructPay, which jointly selects paid workflow cuts and individualized success bonuses to minimize commitment at a fixed worst-case local-reliability target. For an ordered complementary workflow family, we exactly characterize worst risk over approximate coarse correlated equilibria for every nonnegative reward vector. This yields certified optimization and, in a homogeneous regime, matching bonus-commitment bounds, an interior square-root organizational scale and a constant-factor construction. A specified family exhibits an $\Omega(n)$ price of fixed organization. On 300 frozen heterogeneous design instances, StructPay achieves median per-instance commitment savings of 55.5% relative to fully optimized individual rewards on the fixed workflow and 18.1% relative to optimized structure with a common reward. A prespecified 2,700-case continuation shows a strengthening advantage for individualized rewards along the evaluated heterogeneity sequence. In the exact two-agent game, joint design commits 21.4% less while all 16 optimized-contract trajectories across two LLM model families meet the frozen behavioral and reliability targets. Together, these results identify workflow structure and reward allocation as two distinct levers for lowering the commitment required for robust cooperation.

1 Introduction

Reliable multi-service delegation requires organizing work and rewarding hidden, costly effort. A client may observe only a shared verification outcome, while each service bears its own effort cost. Team moral hazard and job design connect incentives to production technology (Holmström, 1982; Holmström & Milgrom, 1991); AI-delegation contracts bring this interaction to model choice, computation and verification (Saig et al., 2024; Sreenivas & Larson, 2026). Two design problems arise when reliability must be achieved with limited commitment.

Problem 1: structural incentive burden. When a downstream check also requires upstream effort, dependence reduces the downstream service's return to working. Larger rewards can compensate, but an independent capability can remove the prerequisite while preserving the deliverable. Even fully optimized individual rewards on a fixed workflow can therefore remain expensive. Figure 1a,d and Table 1 demonstrate this burden: joint design uses a median 55.5% less commitment than the optimized fixed workflow on our 300 controlled instances.

Problem 2: one-size-fits-all rewards. Optimizing the workflow does not settle how to reward its services. Differences in effort costs, verification weights and behavioral tolerances create different incentive needs. A comparator that optimizes both structure and a common reward still leaves an 18.1% median saving for joint design on the same panel. Along a separate prespecified heterogeneity continuation, individualized rewards become increasingly valuable (Figures 1b and 3b,c).

To address both problems, we propose **StructPay**. StructPay jointly chooses workflow dependence and individualized rewards. It minimizes bonus commitment plus restructuring fees at a prescribed reliability target, retaining the required services and returning a workflow, individualized bonuses

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

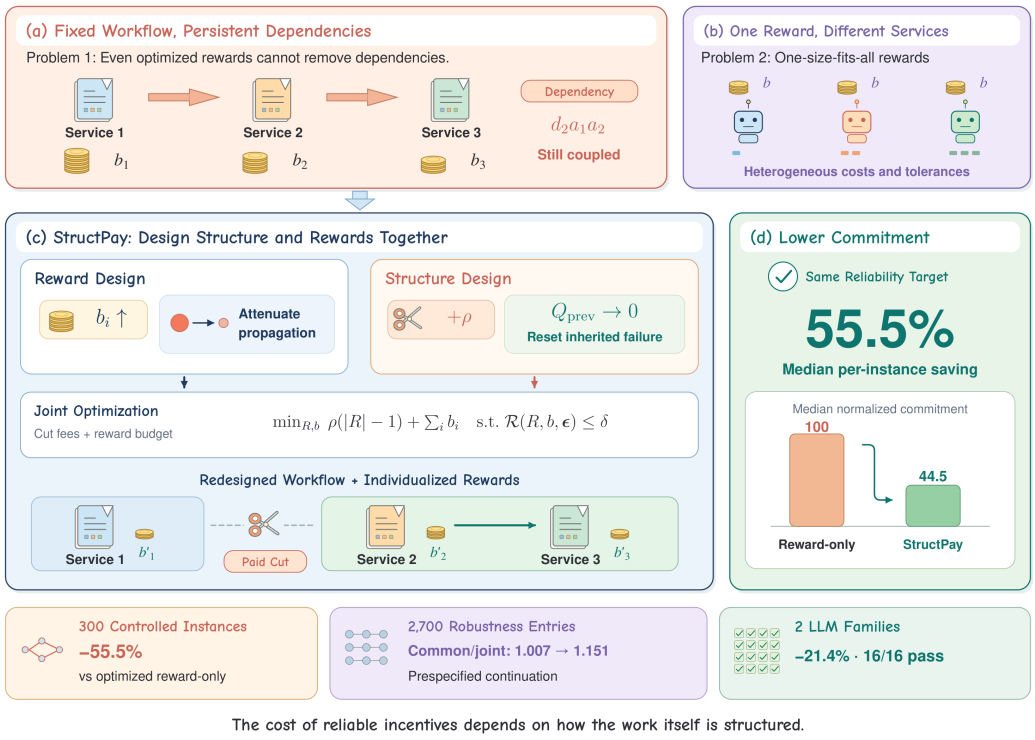

Figure 1: **Two design problems motivate StructPay.** (a,b) Fixed dependencies and a common reward create distinct incentive costs. (c) StructPay jointly designs dependence and individualized rewards. (d) Normalizing each optimized reward-only commitment to 100 gives median paired joint commitment 44.5: a 55.5% saving over 300 instances (Table 1). The lower strip summarizes the separate H2 continuation and E6 test (Figures 3 and 4); 16/16 counts trajectories under both optimized contracts meeting both frozen targets. Service and reward illustrations are schematic.

and a risk certificate. We study an ordered complementary workflow whose prefix structure enables exact robust-risk characterization.

Robust design must meet its target under *every approximate coarse correlated equilibrium* (CCE). This criterion allows correlated play while bounding each service's gain from a fixed unilateral deviation, so reliability does not rely on a favorable equilibrium selection. Its connection to external regret (Freund & Schapire, 1999; Roughgarden, 2015) also provides a measurable behavioral interface: we can test whether real LLM choices meet a tolerance fixed before design.

Our attainable prefix-failure recurrence turns this robust requirement into certified joint optimization. Our contributions are:

- **Exact robust risk.** An attainable $O(n)$ characterization of worst approximate-CCE local failure for every nonnegative reward vector on an ordered complementary workflow.
- **Joint design laws.** Matching commitment bounds, the organizational scale and a constant-factor construction in a homogeneous regime, plus an $\Omega(n)$ separation from optimized fixed organization.
- **Certified evidence.** Controlled gains against both optimized comparators, a 2,700-case heterogeneity continuation and a two-model LLM test: all 16 optimized-contract trajectories meet the frozen targets, with joint commitment 21.4% lower.

2 RELATED WORK

Team production and organization. Team moral hazard (Holmström, 1982) and combinatorial agency (Babaioff et al., 2012) model hidden effort rewarded through observable outcomes. Complementary production can favor unequal rewards even among identical agents (Winter, 2004). Joint technology and payment design is also established: Hatakeyama (2026) optimizes production technology and success bonuses to implement full effort in a unique Nash equilibrium at minimum cost. StructPay instead optimizes worst approximate-CCE local reliability, jointly pricing dependence changes and individualized rewards while retaining all required services.

Contract optimization and imperfect responses. Algorithmic contract theory studies optimal contracts for combinatorial production (Emek & Feldman, 2012) and supermodular rewards (Deo-Campo Vuong et al., 2024). Bacchiocchi et al. (2025) study a single agent's approximate best response. Dütting et al. (2025) give principal-utility guarantees in every CCE for submodular rewards; for binary supermodular rewards, they equate the best-CCE and best-pure-Nash utility benchmarks. StructPay characterizes exact adverse local-failure risk for every nonnegative bonus vector and heterogeneous deviation tolerances. This enables certified commitment minimization at a fixed reliability target in the complementary path family.

AI delegation and self-interested cooperation. Quality contracts (Saig et al., 2024), joint inspection and payment design (Saig et al., 2026), and hidden model and token-effort choices (Sreenivas & Larson, 2026) address AI delegation. Principal-agent reinforcement learning learns sequential contracts and agent policies (Ivanov et al., 2024). Shapley-Coop combines task-time negotiation with Shapley-based post-task redistribution for self-interested LLM cooperation (Hua et al., 2025). StructPay selects paid production-dependence cuts and individualized success bonuses before effort, using an exact worst-equilibrium reliability constraint. Its contribution concerns the commitment required by the jointly designed organization and contract.

Workflow design and LLM behavior. AFlow searches executable agent workflows (Zhang et al., 2025); our cuts are priced changes to a strategic production technology. NegotiationArena evaluates LLM bargaining behavior (Bianchi et al., 2024). Our behavioral test measures external regret and local reliability under frozen contracts, connecting observed choices to the tolerance used in the risk certificate.

3 PROBLEM SETUP

We consider n required services in a public order. A set $R \subseteq [n]$, with $1 \in R$, identifies the roots of contiguous modules. Service i chooses hidden effort $a_i \in \{0, 1\}$ at public cost $c_i a_i$, where $c_i > 0$. Let $d_i > 0$ and $\sum_i d_i = 1$. Success probability and expected utility are

$$p_R(a) = \sum_{i \in R} d_i a_i + \sum_{i \notin R} d_i a_{i-1} a_i, \tag{1}$$

$$u_i(a) = b_i p_R(a) - c_i a_i, \qquad b_i \geq 0. \tag{2}$$

A verification item i is selected with probability d_i. A root item requires its own effort, and another item requires its own and its predecessor's effort. All services receive their respective bonuses upon the same success outcome. Thus p_R measures *local-check reliability*: the probability that the selected verification item succeeds. A cut purchases an independent local capability, turning a nonroot into a root at price $\rho \geq 0$. It preserves the required deliverable, verification weight and other model parameters. For $s = |R|$, the restructuring charge is $K(R) = \rho(s - 1)$.

For a vector $\boldsymbol{\varepsilon} \geq 0$, a distribution μ is an approximate CCE if, for each i and each fixed alternative $r_i \in \{0, 1\}$,

$$\mathbb{E}_\mu[u_i(r_i, a_{-i}) - u_i(a)] \leq \varepsilon_i. \tag{3}$$

The deviations are fixed before observing a recommendation. The tolerance ε_i has utility units; it is not an independent error probability. Define

$$\mathcal{R}(R, b, \boldsymbol{\varepsilon}) = \sup_{\mu \in \mathrm{CCE}_{\boldsymbol{\varepsilon}}(R,b)} \mathbb{E}_\mu[1 - p_R(a)]. \tag{4}$$

Given failure tolerance $\delta \in (0, 1)$, the principal minimizes

$$\mathrm{OPT} = \inf_{R, b \geq 0} \left\{ \rho(|R| - 1) + \sum_i b_i : \mathcal{R}(R, b, \varepsilon) \leq \delta \right\}. \tag{5}$$

The commitment reserves the sum of success bonuses and pays the restructuring fee, both in game utility units. Given a hard cap, the principal can compare the certified commitment directly with that cap.

The base model conditions on participation. With zero outside options, fixed compensation ε_i per service guarantees nonnegative expected utility and is added to commitment while preserving effort incentives (Appendix D.3). Continuous effort with linear costs has the same worst risk by Bernoulli rounding (Appendix B.1).

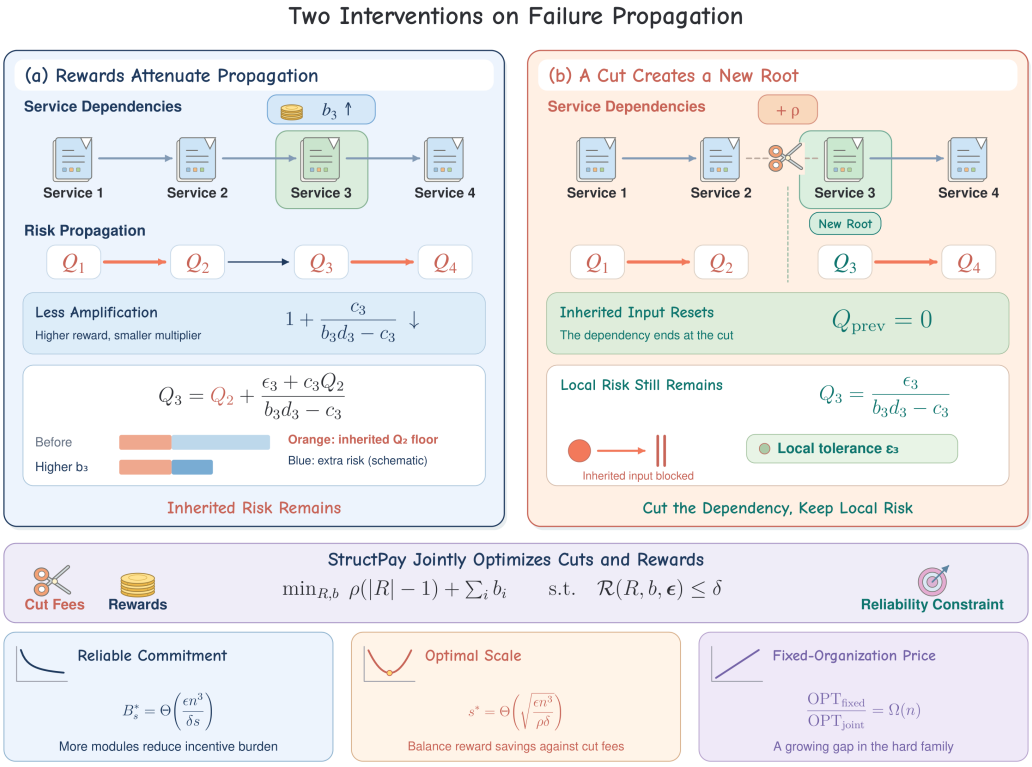

Figure 2: **Rewards attenuate propagation; cuts reset inherited input.** (a) Raising b_3 reduces added risk and amplification above the inherited Q_2 floor. (b) A cut makes service 3 a root with $Q_{\mathrm{prev}} = 0$, retaining its local risk. The displayed recurrences assume positive margins and no saturation; bars and curves are schematic. The footer's bonus-commitment and scale laws assume homogeneous services, $\delta \leq 1/4$ and $\varepsilon \geq c\delta$. The square-root law additionally needs $\rho > 0$ and an interior scale; the $\Omega(n)$ separation holds for the specified family.

4 StructPay: Designing a Workflow and Its Contract

StructPay jointly chooses workflow dependence and individualized rewards. It takes (c, d, ε), target δ and cut price ρ, and returns (R, b) with a reliability certificate. Figure 2 shows its two levers: higher bonuses attenuate added failure and the amplification of inherited risk; a paid cut resets the inherited input. The exact recurrence in Theorem 1 makes their commitment costs directly comparable.

Certified joint optimization. For positive tolerances, StructPay enumerates candidate workflow cuts and paid-prefix patterns, then solves the resulting convex reward subproblems. Feasible contracts give upper bounds; independently verified lower bounds certify the global optimality gap. The

optimizer covers arbitrary individual bonuses and produces certified comparisons for heterogeneous instances. Full certificate construction and pruning appear in Appendix E; exact enumeration may require exponentially many subproblems.

Restricting the solver to the original chain gives the reward-only comparator. Reoptimizing a common bonus on every structure gives the structure-plus-common-reward comparator. Both optimize their entire permitted design class. This isolates the value of choosing structure and allocating rewards jointly.

Fast homogeneous construction. A separate $O(n)$ construction scans balanced module counts and assigns explicit bonuses (Equation (13)). It has a constant-factor guarantee in the homogeneous regime below. Main empirical comparisons use the certified optimizer; the construction's empirical performance is reported separately in Appendix F.

5 THE COST OF RELIABLE COOPERATION

We characterize the worst risk under arbitrary bonuses, then compare paid restructuring with fully optimized rewards on the original workflow.

Exact risk. Let $y_i = b_i d_i - c_i$ be the incentive margin from making service i's own item succeed. Reset $Q_{\mathrm{prev}} = 0$ at each module root, and otherwise use the previous item's value. Define

$$Q_i = \begin{cases} 1, & y_i \le 0, \\ \min\{1, Q_{\mathrm{prev}} + (\varepsilon_i + c_i Q_{\mathrm{prev}})/y_i\}, & y_i > 0. \end{cases} \tag{6}$$

Theorem 1 (Exact worst-case reliability)**.** *For every partition R, positive costs and weights, nonnegative bonus vector and nonnegative tolerance vector,*

$$\mathcal{R}(R, b, \boldsymbol{\varepsilon}) = \sum_i d_i Q_i. \tag{7}$$

The value is computable in $O(n)$ arithmetic operations, and an explicit approximate CCE attains it.

Why exactness matters. An effort deviation bounds prefix failure by Equation (6). Before saturation, larger bonuses reduce both the added slack ε_i/y_i and the inherited multiplier $1 + c_i/y_i$; cuts reset the inherited bound. A matching approximate CCE draws an active prefix independently in each module, with mass $Q_{k+1} - Q_k$ on prefix length k ($Q_0 = 0$, $Q_{m+1} = 1$). It attains all item bounds simultaneously. This attaining equilibrium makes the formula exact, including saturation and zero bonuses (Appendix B). For positive tolerances, fixed-workflow reward optimization reduces to a finite convex family, including unpaid suffixes (Proposition 2).

The commitment law. Assume homogeneous parameters $d_i = 1/n$, $c_i = c > 0$ and $\varepsilon_i = \varepsilon > 0$. Let $B_s^*(\delta, \varepsilon)$ be the infimum bonus commitment over all contiguous partitions with exactly s modules and all bonuses meeting the risk target. Write $A = \varepsilon n^3/\delta$ and $k_\delta = 4(1-\delta)^3/9$.

Theorem 3 (Necessary and achievable commitment)**.** *For $0 < \delta \le 1/4$, $\varepsilon \ge c\delta$ and $1 \le s \le n$,*

$$\frac{3}{16}\frac{\varepsilon n^3}{\delta s} \le B_s^*(\delta, \varepsilon) \le 5\frac{\varepsilon n^3}{\delta s}. \tag{8}$$

The lower bound covers individualized bonuses and unpaid suffixes. The upper bound uses balanced modules: a length-m module receives margin $m(c + \varepsilon/\delta)$. Thus $B_s^* = \Theta(A/s)$ compares an achievable design with *every* contract, without requiring balanced partitions to be exactly optimal. General finite bounds and the proof are in Appendices A and D.

Choosing the organization. Cuts trade the reduction in bonus commitment against $\rho(s-1)$.

Corollary 4 (Optimal organizational scale)**.** *Under the conditions of Equation* (8)*, let s^* minimize $B_s^* + \rho(s-1)$. For $\rho > 0$, set $x = \sqrt{A/\rho}$ and $z = \min\{n, \max\{1, x\}\}$. Then*

$$(k_\delta/6)z \le s^* \le 6z. \tag{9}$$

Thus $s^* = \Theta(\sqrt{\varepsilon n^3/(\rho\delta)})$ when $1 \ll x \ll n$. A sufficient finite interior condition is $6/k_\delta < x < n/6$. Full separation is optimal at $\rho = 0$; the uncut chain is uniquely optimal when $\rho \ge 5A$.

The law identifies the organizational scale at which commitment reduction balances restructuring cost (Appendix D.1). Scanning balanced module counts gives a risk-feasible $O(n)$ construction with $\mathrm{ALG} \leq (80/3)\,\mathrm{OPT}$ under the same homogeneous conditions (Theorem 5). This guarantee is separate from certified exhaustive optimization.

Theorem 6 (Price of fixed organization)**.** *For any $n \geq 16$, set $\delta = 1/10$, $c = 1/(16n^3)$, $\varepsilon = \delta c$ and $\rho = 1/(16n^2)$. The optimum reliable bonus commitment on the unmodified chain is at least $81/4000$, while a feasible fully separated design costs $(3n-1)/(16n^2)$. Both alternatives have feasible total commitment below one. Consequently,*

$$\frac{\mathrm{OPT}_{\mathrm{chain}}}{\mathrm{OPT}_{\mathrm{joint}}} \geq \frac{81n^2}{250(3n-1)} = \Omega(n). \tag{10}$$

The numerator optimizes every reward on the fixed chain; the denominator includes positive cut fees. This existence family establishes that payment optimization alone can be asymptotically more expensive at the same reliability target (Appendix D.2).

6 Controlled design and LLM behavioral evidence

We test structural value after reward optimization (RQ1), individualized-reward value after structure optimization (RQ2), and compatibility of real LLM choices with the designed behavioral regime (RQ3). Commitment is measured in game utility units, not API expenditure.

6.1 Main controlled benchmark

Table 1: **Certified commitment on 300 controlled heterogeneous instances.** All methods meet the same .99 worst-case local-reliability target. Lower commitment ratios are better.

Method	Median C/C_{joint}	Median joint saving
Fixed + optimized common reward	$2.570\times$	61.1%
Fixed + optimized individual rewards	$\mathbf{2.247}\times$	$\mathbf{55.5}\%$
Structure + optimized common reward	$1.221\times$	18.1%
StructPay joint optimizer	$\mathbf{1.000}\times$	—

The primary controlled design panel contains all 300 combinations of four service counts $n \in \{6, 8, 10, 12\}$, five cost patterns, five verification-weight patterns and three tolerance patterns. Costs have mean one, weights sum to one, and tolerances have mean .01. The target $\delta = .01$ and fee $\rho = n^3/25$ were frozen before optimization. All methods meet the same worst-case local reliability of .99. Each comparator is optimized within its design class; enumeration and independent certificates are detailed in Appendix G.

RQ1: structure adds value after reward optimization. Table 1 shows a median per-instance **55.5%** commitment saving against arbitrary individual bonuses fully optimized on the fixed workflow. The structural benefit thus survives asymmetric reward optimization. Each saving is computed as $1 - C_{\mathrm{joint}}/C_{\mathrm{method}}$ before taking the median; it is not a ratio of median costs. The separate homogeneous slice in Figure 3a shows the price boundary: joint design saves at low cut prices and recovers the fixed workflow at high prices.

RQ2: individualized rewards add value after structure optimization. Against a comparator that reoptimizes both the cut mask and common bonus, joint design saves a median **18.1%**. Savings over this comparator exceed 10% in 195/300 cases and 25% in 120/300, certified using unfavorable endpoints. A matched post-hoc diagnostic shows that reoptimizing individual rewards on the common-reward-selected structure reduces its median cost ratio to 1.029, explaining much of the gain, while that structure remains certified suboptimal in $205/300$ instances (Appendix G).

6.2 Robustness across heterogeneity

The H2 continuation varies costs, weights and tolerances jointly from homogeneous to heterogeneous values at five strengths λ. Sixty base patterns, three targets and three prices give 540 entries

per strength: 2,700 entries in total, representing 2,196 unique inputs. Weights satisfy $\min_i d_i > .05$, preserving the unsaturated regime (Appendix H).

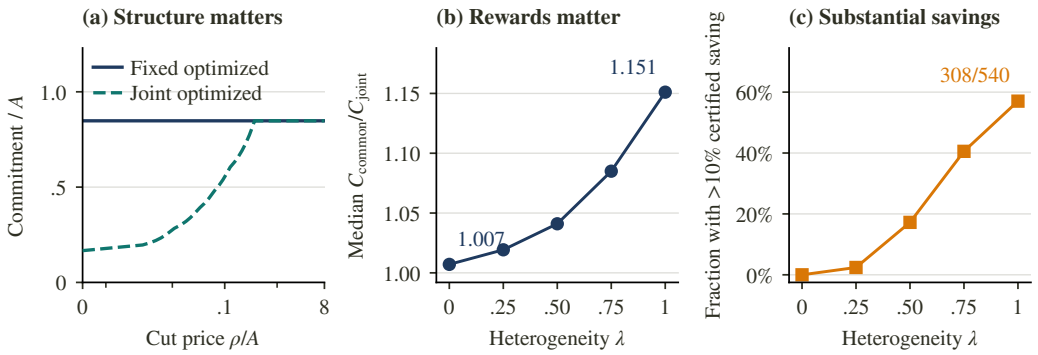

Figure 3: **Structure and rewards provide distinct sources of savings.** (a) Fixed-workflow and joint commitment in a homogeneous slice ($n = 12, c = 1, \delta = \varepsilon = .1$; $A = \varepsilon n^3/\delta$; symlog price axis). (b) Median common-bonus/joint ratio along the heterogeneous continuation. (c) Fraction with certified saving $1 - C_{\text{joint}}/C_{\text{common}} > .10$, using unfavorable endpoints. Each strength in (b,c) pools 540 deterministic entries; these are grid summaries, not population estimates.

The median common-bonus/joint ratio rises from 1.007 to 1.151, while certified savings above 10% increase from $0/540$ to $308/540$ (Figure 3b,c). Table 2 also shows that the fraction of selected common-reward structures retaining a certified loss after individualized reward reoptimization increases. The endpoint advantage strengthens in all nine target/fee settings (Table 3).

Table 2: H2 continuation (540 prespecified entries per row). Ratios compare common-reward and joint commitment. Certified mask loss means the selected common-reward structure remains suboptimal after individualized reward reoptimization: its lower bound exceeds the joint upper bound. Mask fractions describe selected witnesses in this grid.

λ	Median common/joint	$> 10\%$ savings (certified)	Different cut masks	Certified mask loss
0	1.007	0/540	0.056	0.056
.25	1.019	13/540	0.280	0.274
.50	1.041	93/540	0.426	0.415
.75	1.085	219/540	0.530	0.520
1	1.151	308/540	0.565	0.561

Table 3: H2 endpoint sensitivity. Each cell gives the median common/joint commitment ratio from $\lambda = 0$ to $\lambda = 1$ for the same 60 prespecified base patterns; columns are per-cut fees.

δ	$\rho = n^3/100$	$\rho = n^3/25$	$\rho = n^3/5$
.01	$1.000 \to 1.147$	$1.008 \to 1.104$	$1.022 \to 1.075$
.025	$1.002 \to 1.271$	$1.008 \to 1.159$	$1.011 \to 1.121$
.05	$1.002 \to 1.359$	$1.009 \to 1.205$	$1.023 \to 1.167$

6.3 REAL-AGENT BEHAVIORAL COMPATIBILITY

We next test whether real LLM agents satisfy the behavioral tolerance and reliability target used by StructPay.

The exact two-service game fixes $d_i = .5$, $c_i = 1$, $\varepsilon^* = .05$, $\delta^* = .10$ and $\rho = .25$ before model calls. Optimized fixed-chain commitment is 7.9495; joint design cuts the edge and commits 6.25, including its fee: **21.4%** less at the same certified target. This tests exact-risk design, since $\varepsilon^* < c\delta^*$ lies outside the homogeneous scaling regime.

DeepSeek-V4.1-Flash (DeepSeek-AI, 2026) and Qwen3.6-35B-A3B (Qwen Team, 2026) each run four fresh 64-round trajectories per contract. An under-incentivized chain with bonuses $(.5, .5)$ completes the 24-trajectory comparison. Agents see production rules, their contract and private realized feedback; actions commit before feedback. Expected utilities supply offline external regret, and actual choices determine expected local reliability (Appendix I).

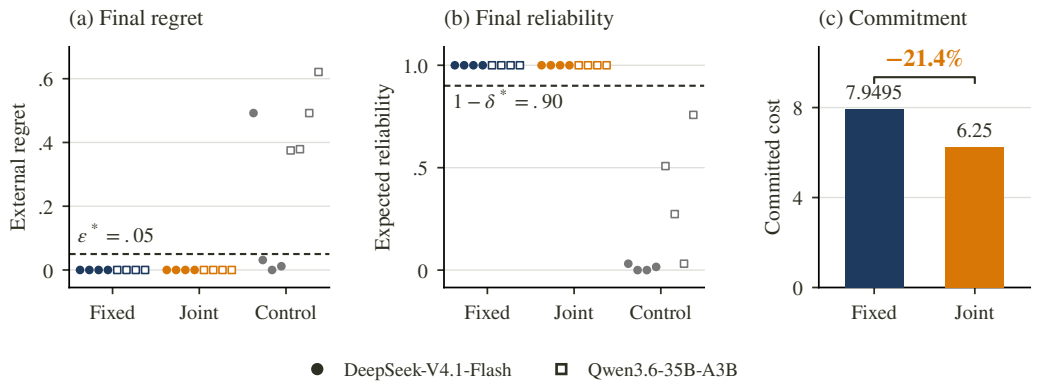

Figure 4: **Lower commitment with compatible LLM behavior.** (a,b) Final maximum-agent external regret and expected local reliability for all 24 fresh 64-round trajectories. Horizontal offsets separate coincident values. (c) Joint commitment includes the cut fee and is 21.4% lower. Table 4 summarizes threshold attainment.

Table 4: **LLM compatibility at the frozen targets.** Counts are fresh trajectories; joint commitment includes its cut fee.

Contract	Commitment	DeepSeek-V4.1-Flash	Qwen3.6-35B-A3B
Fixed optimized	7.9495	4/4 pass	4/4 pass
Joint optimized	**6.25**	**4/4 pass**	**4/4 pass**
Under-incentivized	1.0	0/4 reliable	0/4 reliable

Pass requires both maximum-agent external regret $\leq \varepsilon^* = .05$ and expected local reliability $\geq 1 - \delta^* = .90$. "Reliable" checks the latter target alone.

All **16 optimized-contract trajectories** satisfy both frozen thresholds and choose HH throughout; all eight controls fail reliability. Effort is dominant for both services after the cut and for the upstream service on the optimized chain. Across both model families, the joint contract therefore achieves behavioral compatibility with the specified game at lower commitment.

7 SCOPE AND OUTLOOK

Our exact characterization uses the prefix structure of ordered complementary workflows. Extending joint workflow–reward design to richer dependency graphs, dynamic routing and feedback loops is a natural next step. The guarantee concerns the specified local verification technology rather than correctness of every deliverable, and a real cut must preserve the modeled isolation property. Costs are public and participation is conditioned on; private-cost elicitation and richer outside options require additional modeling. The exact heterogeneous optimizer can require exponentially many subproblems, motivating scalable approximations. Our deterministic design grids provide controlled comparisons; jointly varying heterogeneity parameters does not isolate their individual causal effects or imply universal monotonicity. The finite LLM trajectories establish behavioral compatibility in the modeled game. Additional behavioral and application diagnostics, including positive and negative outcomes, are retained in the internal research archive.

8 CONCLUSION

StructPay jointly chooses workflow dependence and individualized rewards. Its exact worst-risk characterization yields reliable-commitment and organizational-scale laws, explaining how rewards attenuate failure propagation while cuts reset its inherited component. Certified experiments show median per-instance commitment savings of 55.5% over fully optimized reward-only design and 18.1% over optimized structure with a common reward. Real LLM trajectories satisfy the designed tolerance and reliability target under a joint contract with 21.4% lower commitment. These results establish workflow structure and reward allocation as distinct design levers for robust cooperation. Future work will extend the framework to richer workflow topologies and scalable heterogeneous optimization.

REPRODUCIBILITY STATEMENT

Complete proofs, controlled protocols and all E6 trajectories accompany the manuscript. Code and compact records are available at https://anonymous.4open.science/r/anonymousgit-BBF2/. The source bundle includes figures, plotted data and source hashes; the code provides optimization, benchmark generation, behavioral replay and opt-in live evaluation. Exhaustive certificates, raw calls and additional diagnostics remain in the internal research archive. Grid entries are deterministic design cases; LLM units are fresh trajectories. Recorded inputs reproduce mathematical and scoring outputs, while fresh generations remain stochastic. Reproduction scope and instructions appear in Appendix J.

AI USE STATEMENT

AI assistance was used only to polish the writing of this manuscript.

REFERENCES

Moshe Babaioff, Michal Feldman, Noam Nisan, and Eyal Winter. Combinatorial agency. *Journal of Economic Theory*, 147(3):999–1034, 2012. doi: 10.1016/j.jet.2012.01.010.

Francesco Bacchiocchi, Jiarui Gan, Matteo Castiglioni, Alberto Marchesi, and Nicola Gatti. Contract design under approximate best responses. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 2203–2220, 2025. URL https://proceedings.mlr.press/v267/bacchiocchi25a.html.

Federico Bianchi, Patrick John Chia, Mert Yuksekgonul, Jacopo Tagliabue, Dan Jurafsky, and James Zou. How well can LLMs negotiate? NegotiationArena platform and analysis. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 3935–3951. PMLR, 2024. URL https://proceedings.mlr.press/v235/bianchi24a.html.

DeepSeek-AI. DeepSeek-V4.1-Flash: Pushing the limits of KV cache compression. arXiv:2609.19969, version 1, 2026. URL https://arxiv.org/abs/2609.19969v1.

Ramiro Deo-Campo Vuong, Shaddin Dughmi, Neel Patel, and Aditya Prasad. On supermodular contracts and dense subgraphs. In *Proceedings of the 2024 Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 109–132, 2024. doi: 10.1137/1.9781611977912.6. URL https://arxiv.org/abs/2308.07473.

Paul Dütting, Tomer Ezra, Michal Feldman, and Thomas Kesselheim. Black-box lifting and robustness theorems for multi-agent contracts. arXiv:2511.19358, version 1, 2025. URL https://arxiv.org/abs/2511.19358v1.

Yuval Emek and Michal Feldman. Computing optimal contracts in combinatorial agencies. *Theoretical Computer Science*, 452:56–74, 2012. doi: 10.1016/j.tcs.2012.05.018. URL https://tik-db.ee.ethz.ch/file/79d3e25e3233bd72c198ec0afe8bd39b/.

Yoav Freund and Robert E. Schapire. Adaptive game playing using multiplicative weights. *Games and Economic Behavior*, 29(1–2):79–103, 1999. doi: 10.1006/game.1999.0738. URL https://cseweb.ucsd.edu/~yfreund/papers/games_long.pdf.

Tetsutaro Hatakeyama. Incentives and discrimination: A functional optimization approach. *Economics Letters*, 262:112889, 2026. doi: 10.1016/j.econlet.2026.112889. URL https://ies.keio.ac.jp/upload/DP2025-003_EN.pdf. Working-paper version, Keio DP2025-003, consulted for the model and proofs.

Bengt Holmström. Moral hazard in teams. *The Bell Journal of Economics*, 13(2):324–340, 1982. doi: 10.2307/3003457. URL https://people.duke.edu/~qc2/BA532/1982%20Rand%20Holmstrom%20team.pdf.

Bengt Holmström and Paul Milgrom. Multitask principal–agent analyses: Incentive contracts, asset ownership, and job design. *The Journal of Law, Economics, and Organization*, 7(special issue): 24–52, 1991. doi: 10.1093/jleo/7.special_issue.24. URL https://www.stanford.edu/~milgrom/publishedarticles/Multitask%20Principal%20Agent.pdf.

Yun Hua, Haosheng Chen, Shiqin Wang, Wenhao Li, Xiangfeng Wang, and Jun Luo. ShapleyCoop: Credit assignment for emergent cooperation in self-interested LLM agents. In *Advances in Neural Information Processing Systems*, volume 38, pp. 88675–88702. Curran Associates, Inc., 2025. doi: 10.52202/085713-2965.

Dima Ivanov, Paul Dütting, Inbal Talgam-Cohen, Tonghan Wang, and David C. Parkes. Principal-agent reinforcement learning: Orchestrating AI agents with contracts. arXiv:2407.18074, version 2, 2024. URL https://arxiv.org/abs/2407.18074v2.

Qwen Team. Qwen3.6-35B-A3B: Agentic coding power, now open to all. Official model release, April 2026. URL https://qwen.ai/blog?id=qwen3.6-35b-a3b.

Tim Roughgarden. Intrinsic robustness of the price of anarchy. *Journal of the ACM*, 62(5):1–42, 2015. doi: 10.1145/2806883. URL https://theory.stanford.edu/~tim/papers/robust.pdf.

Eden Saig, Ohad Einav, and Inbal Talgam-Cohen. Incentivizing quality text generation via statistical contracts. In *Advances in Neural Information Processing Systems*, volume 37, 2024. URL https://papers.neurips.cc/paper_files/paper/2024/file/5b93ce41ac6de2bf9aca7e4ba5ba01d5-Paper-Conference.pdf.

Eden Saig, Tamar Garbuz, Ariel D. Procaccia, Inbal Talgam-Cohen, and Jamie Tucker-Foltz. Adaptive contracts for cost-effective AI delegation. arXiv:2603.17212, version 2, 2026. URL https://arxiv.org/abs/2603.17212v2.

Nanda Kishore Sreenivas and Kate Larson. Contracting for LLM delegation: Moral hazard in technology and effort choice. arXiv:2608.18232, version 1, 2026. URL https://arxiv.org/abs/2608.18232v1.

Eyal Winter. Incentives and discrimination. *American Economic Review*, 94(3):764–773, 2004. doi: 10.1257/0002828041464434.

Jiayi Zhang, Jinyu Xiang, Zhaoyang Yu, Fengwei Teng, XiongHui Chen, Jiaqi Chen, Mingchen Zhuge, Xin Cheng, Sirui Hong, Jinlin Wang, Bingnan Zheng, Bang Liu, Yuyu Luo, and Chenglin Wu. AFlow: Automating agentic workflow generation. In *International Conference on Learning Representations*, pp. 34040–34077, 2025. URL https://openreview.net/pdf?id=z5uVAKwmjf.

A FORMAL STATEMENTS AND PROOF ROADMAP

The main text states exact risk (Theorem 1), the commitment law (Theorem 3), organizational scale (Corollary 4), and the price of fixed organization (Theorem 6). Here we supply the complementary fixed-workflow reduction and constructive guarantee, retaining the numbering of all six results. Appendix B proves exact risk and the continuous-effort extension; Appendix C proves the reward reduction; Appendix D proves the commitment, scale, and separation results. Appendix E justifies the independently checked numerical bounds.

Proposition 2 (Exact reward optimization on a fixed workflow)**.** *For a fixed partition and $\varepsilon_i > 0$, optimization over all nonnegative bonus vectors reduces exactly to a finite family of convex programs indexed by one paid prefix per module. The family may contain exponentially many patterns.*

Finite commitment bounds. Assume $d_i = 1/n$, $c_i = c > 0$ and $\varepsilon_i = \varepsilon > 0$. Let $B_s^*(\delta, \varepsilon)$ denote the infimum bonus commitment over *all* contiguous partitions with exactly s modules and *all* bonuses satisfying the risk target. Put $A = \varepsilon n^3/\delta$ and $k_\delta = 4(1-\delta)^3/9$.

Theorem 3 follows from the following more general bounds. For $0 < \delta < 1$ and $1 \le s \le n$,

$$B_s^* \ge k_\delta A/s, \qquad B_s^* \ge (1-\delta)n^2 c. \tag{11}$$

For a balanced partition with lengths $m_1, \dots, m_s$, a feasible contract has commitment

$$U_s^{\text{bonus}} = n^2 c + n(c + \varepsilon/\delta) \sum_{j=1}^{s} m_j^2. \tag{12}$$

When $0 < \delta \le 1/4$ and $\varepsilon \ge c\delta$, these bounds give Equation (8).

Constructive design. The same bounds give an implementable homogeneous contract. For each $s = 1, \dots, n$, write $n = qs + t$ with $0 \le t < s$ and compute

$$U_s = \rho(s-1) + n^2 c + n(c + \varepsilon/\delta)\big[(s-t)q^2 + t(q+1)^2\big]. \tag{13}$$

Select the smallest U_s, form the balanced partition, and give each member of a length-m module bonus $n[c + m(c + \varepsilon/\delta)]$.

Theorem 5 (Constant-factor joint design)**.** *Under the parameter conditions of Equation* (8)*, the algorithm returns a risk-feasible contract in $O(n)$ arithmetic operations, with*

$$\text{ALG} \le (5/k_\delta)\,\text{OPT} \le (80/3)\,\text{OPT}\,. \tag{14}$$

This homogeneous approximation has a distinct computational role from the exhaustive small-instance solver: exact risk evaluation is linear in n, whereas general reward and structure optimization can require exponentially many convex subproblems. The implementation and its empirical comparison with optimized alternatives are retained in Appendix F.

B EXACT-RISK PROOF AND EXTENSIONS

Within one module, relabel its positions $1, \dots, m$, retaining their original parameters. Set $I_0 = 0$, and let I_i indicate that at least one position up to i did not exert effort. Write $z_i = \mathbb{E}I_i$. On the event $I_{i-1} = 0, a_i = 0$, switching i to one increases its own verification term by d_i and can only add a nonnegative downstream term. Its utility gain is at least $y_i = b_i d_i - c_i$. On the event $I_{i-1} = 1$, the same deviation gains at least $-c_i$; if i already exerts effort the deviation gain is zero. Elsewhere the gain is zero. Since $\Pr(I_{i-1} = 0, a_i = 0) = z_i - z_{i-1}$,

$$\mathbb{E}[u_i(1, a_{-i}) - u_i(a)] \geq y_i(z_i - z_{i-1}) - c_i z_{i-1}. \tag{15}$$

If $y_i > 0$, the CCE inequality implies

$$z_i \leq \min\left\{1, z_{i-1} + \frac{\varepsilon_i + c_i z_{i-1}}{y_i}\right\}. \tag{16}$$

The right side is nondecreasing in z_{i-1}, giving $z_i \leq Q_i$ by induction. For $y_i \leq 0$, use $z_i \leq 1 = Q_i$. A failed local item implies a failed prefix, so $\mathbb{E}[1 - p_R] \leq \sum_i d_i Q_i$ across modules.

For the reverse inequality, the sequence Q_i is nondecreasing within a module. Let $Q_0 = 0, Q_{m+1} = 1$ and assign mass $Q_{k+1} - Q_k$ to the action with precisely its first k positions active, $0 \leq k \leq m$. The masses are nonnegative and sum to one. A local item at i succeeds exactly when $K \geq i$, so its failure probability is Q_i.

Fix position i. If $K = i - 1$, switching on gains y_i: its predecessor is on (or it is a root), but its successor is off (or in another module). If $K < i - 1$, neither the needed predecessor nor successor is active, and the gain is $-c_i$. If $K \geq i$, the service is already active. Therefore the expected gain from switching on is exactly

$$y_i(Q_i - Q_{i-1}) - c_i Q_{i-1}. \tag{17}$$

For positive y_i, the truncated recurrence ensures that this is at most ε_i. For nonpositive y_i, it is nonpositive.

For switching off, any position that is active with positive probability has $Q_i < 1$, hence $y_i > 0$. Whenever active, its predecessor is active unless it is a root. Turning off thus loses at least $b_i d_i$ and saves only c_i, a strictly negative gain. A position with $Q_i = 1$ is never active and has zero off-deviation gain. This verifies the second deviation, including all saturated positions. Take independent prefix draws in different modules. Since terms in other modules cancel under a unilateral deviation, all constraints still hold, and every item's failure probability is attained simultaneously. This proves Theorem 1.

The lower-bound distribution is an adversarial witness, not a claim about a typical realized workflow. The prefix shape is used to solve the worst-case problem; no LLM trajectory is assumed to follow it. Figure 5 illustrates how this witness attains the recurrence before and after a cut.

B.1 CONTINUOUS EFFORT AND COVERAGE

The finite action space is $\{0, 1\}^n$, and a CCE distribution may correlate all services. The repeated interaction, observation history, or learning procedure that might produce a distribution is not specified by the model. A tolerance of zero gives the usual CCE constraints. Nonnegative tolerances enlarge that set, so the worst risk is nondecreasing in each tolerance.

Lemma 7 (Rounding equivalence)**.** *Replace binary effort by $q_i \in [0, 1]$ in Equations* (1)–(2)*, retaining linear cost. For every structure, bonus vector and tolerance vector, the continuous and binary games have identical worst CCE risk.*

Proof. Take any distribution on continuous vectors q. Conditional on q, round coordinates independently: $a_i \sim \text{Bernoulli}(q_i)$. Since every production term is multilinear and involves distinct coordinates, $\mathbb{E}[p_R(a) \mid q] = p_R(q)$. Also $\mathbb{E}[c_i a_i \mid q] = c_i q_i$. The same equalities hold after fixing coordinate i to either endpoint $r_i \in \{0, 1\}$. Hence expected endpoint-deviation gains and expected risk are preserved. Own-action utility is affine, so every fixed interior deviation is a convex combination of the two endpoint deviations; endpoint inequalities suffice. Rounding any continuous CCE therefore gives a binary CCE of identical risk. Conversely a binary distribution embeds into the continuous action space, where interior deviations add no constraint. Both worst risks are equal. The

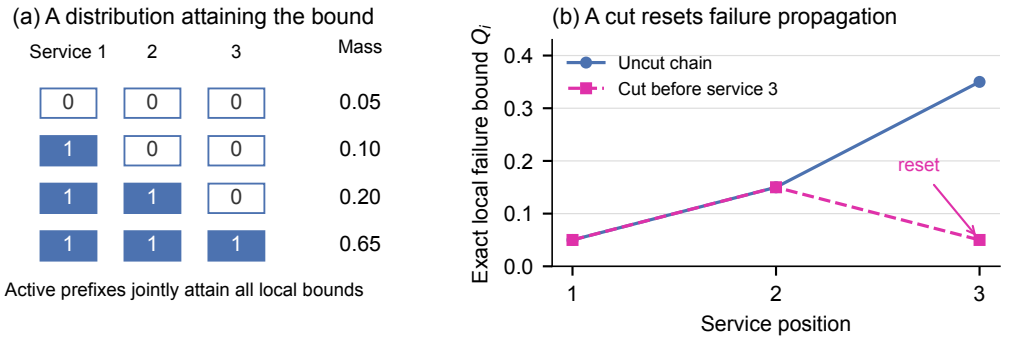

Figure 5: A matching worst equilibrium explains exactness. For three services with $c_i = 1$, $d_i = 1/3$, $b_i = 6$ and $\varepsilon_i = 1/20$, the active-prefix distribution attains $Q = (.05, .15, .35)$. Cutting before service 3 resets inherited failure, giving $Q = (.05, .15, .05)$ and changing risk from $11/60$ to $1/12$. Bonuses are held fixed; the cut additionally costs ρ. This is a recurrence illustration, not an equal-cost comparison.

argument preserves all bonuses, so it also preserves the minimum commitments under any fixed risk target. □

The affine-cost restriction is essential to this proof: conditional rounding has $\mathbb{E}[a_i^2 \mid q_i] = q_i$, not q_i^2. We do not identify an extra call with binary zero-to-one effort in the empirical study.

Local-to-global coverage. Let F denote whole-task failure and F_i failure of monitored item i, on the same outcome space. Assume $F \subseteq \bigcup_i F_i$ and that the local-check risk is $\mathbb{E}[1 - p_R] = \mathbb{E}[\sum_i d_i \mathbf{1}\{F_i\}]$. On F, this sum is at least $d_{\min} > 0$, so $\mathbf{1}\{F\} \leq \sum_i d_i \mathbf{1}\{F_i\}/d_{\min}$ pointwise. Taking expectations and capping probability at one proves the bound without independence. The identity is exact for the monitored production terms in Equation (1); arbitrary proxy metrics need not satisfy it. The bound is sharp under coverage alone: allow only a minimum-weight monitored item to fail, with probability $\delta/d_{\min}$ when $\delta \leq d_{\min}$, and always when $\delta \geq d_{\min}$. Coverage must hold under each structure being compared.

C FIXED-WORKFLOW REWARD OPTIMIZATION

If $Q_i = 1$ in a module, every subsequent recurrence value equals one. Replacing all bonuses from the first such position onward by zero leaves every recurrence value unchanged and weakly reduces cost. Thus one may restrict the search to a paid prefix of each module without excluding an optimal value or an infimizing sequence.

Suppose all $\varepsilon_i > 0$. Enumerate the paid prefix lengths in a fixed partition. For retained positions let $y_i > 0$ and define the untruncated recursion

$$\phi_i = \phi_{\text{prev}} + \frac{\varepsilon_i + c_i \phi_{\text{prev}}}{y_i}, \qquad \phi_{\text{prev}} = 0 \text{ at a root.} \tag{18}$$

Constrain each paid prefix's terminal $\phi_i \leq 1$ and solve

$$\inf_{y_i > 0} \quad \sum_{i \in P} (c_i + y_i)/d_i, \tag{19}$$

$$\text{subject to} \quad \sum_{i \in P} d_i \phi_i + \sum_{i \notin P} d_i \leq \delta. \tag{20}$$

Here P is the union of the selected prefixes. Each ϕ_i is a sum of inverse monomials with nonnegative coefficients. A monomial $f(y) = \prod_j y_j^{-\alpha_j}$ with $\alpha_j \geq 0$ is convex on the positive orthant: for any vector v,

$$v^\top \nabla^2 f(y) v = f(y) \left[\left(\sum_j \alpha_j v_j / y_j \right)^2 + \sum_j \alpha_j (v_j / y_j)^2 \right] \geq 0. \tag{21}$$

Thus each subproblem is convex. Any original contract can be prefix-reduced to one such pattern without increasing cost. Conversely every feasible subproblem supplies a contract whose exact risk meets the target; terminal equality can be handled by dropping saturated payments. Taking the infimum over all patterns gives the true fixed-structure optimum. The pattern count can be exponential in the number of modules. These facts do not provide a polynomial-time global solver for heterogeneous instances.

D COMMITMENT, SCALE AND SEPARATION PROOFS

Consider any feasible partition and bonus vector in the homogeneous model. Let k_j be the count of unsaturated positions in module j. Saturation persists along a module, so these positions form a prefix. Set $K = \sum_j k_j$. The risk constraint implies $n - K \le n\delta$, hence $K \ge (1-\delta)n$. Every unsaturated position has $y_i > 0$ and satisfies

$$Q_i \ge Q_{i-1} + \varepsilon / y_i. \tag{22}$$

Summing across prefixes gives

$$n\delta \ge \sum_i Q_i \ge \varepsilon \sum_{j=1}^{s} \sum_{i=1}^{k_j} (k_j - i + 1)/y_{j,i}. \tag{23}$$

Cauchy–Schwarz yields

$$\left(\sum_{j,i} y_{j,i} \right) \left(\sum_{j,i} \frac{k_j - i + 1}{y_{j,i}} \right) \ge \left(\sum_j \sum_{h=1}^{k_j} \sqrt{h} \right)^2. \tag{24}$$

Since $b_i = n(c+y_i)$ on these positions, discarding the nonnegative nc terms gives the first inequality in Equation (25). Integral comparison gives $\sum_{h=1}^{k} \sqrt{h} \ge (2/3)k^{3/2}$; convexity gives $\sum_j k_j^{3/2} \ge K^{3/2}/\sqrt{s}$. Therefore $\sum_i b_i \ge 4\varepsilon K^3/(9\delta s) \ge k_\delta A/s$.

$$\sum_i b_i \ge \frac{\varepsilon}{\delta} \left(\sum_j \sum_{h=1}^{k_j} \sqrt{h} \right)^2 \ge \frac{4\varepsilon K^3}{9\delta s}. \tag{25}$$

Retaining only the nc terms gives $\sum_i b_i \ge nKc \ge (1-\delta)n^2 c$. This argument permits arbitrary bonus vectors and zero-payment suffixes.

For the upper bound, take a module of length m and a common margin $y = m(c + \varepsilon/\delta)$. Its untruncated terminal value is

$$Q_m^{\text{untruncated}} = \frac{\varepsilon}{c} \left[(1 + c/y)^m - 1 \right]. \tag{26}$$

For $z \ge 0, mz < 1$, the binomial expansion implies $(1+z)^m \le \sum_{k \ge 0} (mz)^k = 1/(1-mz)$. Thus

$$Q_m^{\text{untruncated}} \le \frac{m\varepsilon}{y - mc} = \delta. \tag{27}$$

All preceding values are at most this terminal value. With weights $1/n$, the whole workflow risk is at most δ. Summing $n(c+y)$ over all services yields Equation (12).

For balanced lengths, $\sum_j m_j^2 \le n\lceil n/s \rceil \le 2n^2/s$. If $\varepsilon \ge c\delta$, then $c \le \varepsilon/\delta$ and

$$U_s^{\text{bonus}} \le \varepsilon n^2/\delta + 4\varepsilon n^3/(\delta s) \le 5\varepsilon n^3/(\delta s). \tag{28}$$

For $\delta \le 1/4$, $k_\delta \ge 3/16$. This proves Theorem 3. It does not identify the exact constant or the minimizing partition among all contracts.

D.1 OPTIMAL ORGANIZATIONAL SCALE

Proof. The fixed-reliability theorem gives $k_\delta A/s \le B_s^* \le 5A/s$. Write $H_s = \rho s + B_s^*$, so minimizing H_s is equivalent to minimizing total cost. If $1 \le x \le n$, choose $t = \lceil x \rceil$. With $r = t/x \in [1,2]$,

$$H_{s^*} \le \rho t + 5A/t = (r + 5/r)\sqrt{\rho A} \le 6\sqrt{\rho A}.$$

The last inequality follows because the convex function $r + 5/r$ has maximum 6 on $[1,2]$. Separately bounding H_{s^*} below by ρs^* and by $k_\delta A/s^*$ proves $(k_\delta/6)x \le s^* \le 6x$. The arithmetic–geometric mean inequality gives $H_{s^*} \ge 2\sqrt{k_\delta \rho A}$.

If $x < 1$, comparison with $t = 1$ gives $H_{s^*} \leq \rho + 5A < 6\rho$, hence $s^* < 6$; its lower bound follows from $s^* \geq 1$. If $x > n$, comparison with $t = n$ gives $H_{s^*} \leq \rho n + 5A/n < 6A/n$, hence $s^* > (k_\delta/6)n$; its upper bound follows from $s^* \leq n$. These three cases establish the clipped bound. Since $k_\delta \geq 3/16$, its lower constant is at least $1/32$.

For $\rho = 0$, making every position a root weakly decreases each exact recurrence value for any unchanged bonus vector, so full separation has minimum bonus cost. For $\rho \geq 5A$ and $s \geq 2$,

$$\rho(s-1) + B_s^* \geq 5A + k_\delta A/s > 5A \geq B_1^*,$$

which proves the uncut endpoint. The finite interior assertion follows directly from the two strict bounds. □

D.2 Joint design and fixed-organization separation

Each value U_s in Equation (13) is evaluated in constant arithmetic time once $n, s, c, \varepsilon, \delta, \rho$ are specified. Enumerating s and outputting the chosen bonuses costs $O(n)$ arithmetic operations. This is not a bit-complexity guarantee for arbitrarily encoded real inputs. For any feasible competitor with r modules and total cost C, Equation (11) implies $C \geq \rho(r-1) + k_\delta A/r$. Our algorithm has cost no greater than $U_r \leq \rho(r-1) + 5A/r \leq (5/k_\delta)C$, since $5/k_\delta \geq 1$. Taking the infimum over competitors proves Theorem 5.

For Theorem 6, $\delta = 1/10$ gives $k_\delta = 81/250$ and $A = 1/16$. Equation (11) on the uncut chain gives $B_1^* \geq 81/4000$. Fully separate the services and choose margin $y_i = \varepsilon/\delta = c$. Each root has risk $Q_i = \varepsilon/c = \delta$ and bonus $2nc$. Including $n-1$ cuts gives

$$C_{\text{split}} = 2n^2c + \rho(n-1) = \frac{3n-1}{16n^2}. \tag{29}$$

The chain construction of Equation (12) costs $(2+1/n)/16 < 1$ and meets the same risk target. The separated design also costs less than one, and for $n \geq 16$ its cost is strictly below the chain lower bound. Dividing a lower bound on the chain optimum by an upper bound on the joint optimum gives Equation (10). Full separation need not be the exact joint optimum.

The scaling with n is not an artifact of a vanishing cost unit. Multiplying costs, tolerances, bonuses and restructuring charges together by $16n^3$ preserves every normalized deviation inequality and risk. In the resulting family $c = 1$, $\varepsilon = .1$, and each cut costs n.

As an additional ideal-behavior comparison, if every $y_i > 0$, the first service in each module must work in any exact CCE, and induction forces all subsequent services to work. If the first nonpositive margin in a module occurs at i, a pure equilibrium can stop just before that position; earlier positive margins support the prefix, and later services gain nothing by unilaterally restoring a missing adjacent link. Thus the infimum commitment for unique full implementation at zero tolerance is n^2c for every partition. With positive cut prices, optimizing only that ideal implementation criterion selects the uncut chain. This observation concerns an infimum at strict incentive boundaries; it is not used as a substitute for the optimized reliable comparator in Theorem 6.

D.3 Participation and uniform-error adaptation

For the binary or continuous base model, deviating to zero effort has nonnegative expected utility because production and bonuses are nonnegative. Therefore Equation (3) implies $\mathbb{E}u_i \geq -\varepsilon_i$. Adding a fixed, action-independent payment of ε_i guarantees nonnegative expected utility for a zero outside option and leaves the CCE set unchanged. The principal must add $\sum_i \varepsilon_i$ to its commitment. Both the chain and the joint construction in the separation family can receive this compensation, and the cost-ratio growth remains linear. This is not a theorem about participation under arbitrary private outside options.

For a general finite production game, write the true and estimated production functions as $p, \hat{p}$, and the respective cost functions as $C_i, \hat{C}_i$. Suppose $|p(a) - \hat{p}(a)| \leq \kappa$ for all profiles and

$$|[C_i(r_i, a_{-i}) - C_i(a)] - [\hat{C}_i(r_i, a_{-i}) - \hat{C}_i(a)]| \leq \nu_i \tag{30}$$

for every fixed deviation and profile. Then the difference between the true and estimated deviation gains is at most $2b_i\kappa + \nu_i$. Consequently a true ε_i-CCE belongs to the estimated game's $(\varepsilon_i + 2b_i\kappa + \nu_i)$-CCE set; evaluating the latter's worst failure and adding κ gives a true-risk upper bound when the uniform error assumptions hold.

E CERTIFIED JOINT OPTIMIZER

Homogeneous reduction. Write u for the number of unpaid positions and $K = n - u$ for the number of paid positions after prefix reduction. Feasibility requires $u < n\delta$, since paid positions have strictly positive recurrence values. For homogeneous d_i, c_i, ε_i, the convex bonus problem depends only on u and the multiset of positive paid-prefix lengths: changing module order or assigning the same unpaid positions to different module suffixes changes neither the objective nor any paid recurrence. A completely unpaid module can be removed after reordering modules and its positions appended to another module's unpaid suffix. This leaves all paid recurrences and total unpaid risk unchanged and weakly lowers restructuring cost for $\rho \geq 0$. Some paid module exists because $\delta < 1$.

Consequently it suffices to enumerate every integer u with $0 \leq u < n\delta$ and every integer partition $(k_1, \ldots, k_s)$ of $n - u$ into positive parts. Assigning the u unpaid positions to the last module realizes each enumerated pattern as a contiguous partition of the original n homogeneous positions. This proves equivalence to the original joint optimization, rather than merely restriction to balanced modules. For each pattern, use $c = 1$ and solve $\min \sum_i y_i$ subject to $\sum_i \phi_i \leq n\delta - u$ and the terminal constraints. Its bonus cost is $n(K + \sum_i y_i)$. Reuse its result at each price after adding $\rho(s - 1)$.

Global lower-bound certificate. For a pattern, write the vector of convex constraints as $g(y) \leq 0$, its margin objective as $f(y) = \sum_i y_i$, and its Lagrangian as $L(y; \lambda) = f(y) + \lambda^\top g(y)$ for any $\lambda \geq 0$. Let y^+ be a verified feasible point and $U = f(y^+)$. Every optimum lies in the box $0 < y_i \leq U$. At any positive candidate $\bar{y}$, convexity gives

$$L(v; \lambda) \geq L(\bar{y}; \lambda) + \nabla L(\bar{y}; \lambda)^\top (v - \bar{y}).$$

For every feasible v in that box, $f(v) \geq L(v; \lambda)$. Thus, with $h = \nabla L(\bar{y}; \lambda)$,

$$\max\left\{0,\ L(\bar{y}; \lambda) + \sum_i \min\{-h_i \bar{y}_i,\ h_i(U - \bar{y}_i)\}\right\} \tag{31}$$

is a global lower bound on the optimal margin objective. No stationarity or solver success assumption is needed: a poor candidate only widens this bound. Adding the constant effort-cost term and restructuring cost gives a pattern-level cost interval. The minimum of all pattern lower bounds and the minimum of all pattern upper bounds enclose the joint optimum.

Independent arithmetic. SLSQP generates candidate points and multipliers. It is not used as an optimality certificate. An independent standard-library checker interprets their saved decimal values as exact rational numbers. It evaluates every recurrence, derivative, feasibility inequality and displayed lower bound using rational arithmetic, then rounds exported interval endpoints outwards. All 1,560 prefix certificates pass. The maximum relative pattern interval is 5.39×10^{-7}, and the maximum relative joint interval is 3.83×10^{-7}. Solver diagnostics and amendments remain in the internal research archive; candidate status is never an optimality certificate.

Completeness and discrete optimizer checks. A second enumeration uses every cut mask of the original ordered chain and every paid-prefix choice, independently of the integer-partition generator. For $n = 2, \ldots, 8$ and $\delta \in \{.1, .2, .25\}$, all 21 canonical-set comparisons agree exactly. To recover fixed-s optima, a pattern with r nonempty paid prefixes and u unpaid positions can include h completely unpaid modules for every integer $0 \leq h \leq u$, giving $s = r + h$ at unchanged bonus cost. Minimizing over these eligible patterns provides fixed-s intervals, including patterns with completely unpaid modules. In all 240 joint-grid instances, the upper cost bound for the reported module count is strictly below the lower cost bound for every other module count. Thus each reported module count is uniquely optimal, independently of the numerical solver's stopping status.

Interval convention and heterogeneous inputs. For certified bounds $L \leq \mathrm{OPT} \leq U$, we report relative width $(U - L)/U$ for the positive commitments reported here. Feasible implemented costs and the associated outward optimum intervals are both retained; no solver status substitutes for these bounds. For heterogeneous positive weights the margin objective is $\sum_i y_i/d_i$, and a feasible margin cost U restricts relevant points to $0 < y_i \leq d_i U$. The same supporting-affine argument applies on this weighted box. Appendix G gives the full weighted bound and the strict-target condition that makes every ordered cut mask unsaturated.

F HOMOGENEOUS CONTROLLED CHECKS

Frozen comparison. The frozen homogeneous grid uses: $n \in \{4, 6, 8, 10, 12\}$, $c = 1$, $\delta \in \{.1, .2\}$, $\varepsilon/(c\delta) \in \{1, 2, 4\}$, and $\rho/A \in \{0, .0025, .01, .04, .16, .64, 2, 8\}$, where $A = \varepsilon n^3/\delta$. All 240 instances and all six variants are retained. The protocol and its hash were saved before computing the ablations. The prespecified absolute examples fix $n = 12$, $c = 1$, $\varepsilon = \delta$, and $\rho/A = .04$, at both reliability targets (Table 5).

Operational definitions. The *No mechanism* baseline uses the fixed chain and zero bonuses, whose exact worst risk is one. Equal reward minimizes a common scalar $b_i = b$ on that chain. Reward-only optimizes all nonnegative bonuses on the fixed chain and is the primary comparator. Structure plus equal reward enumerates every integer partition of n and optimizes a common scalar separately for each. Homogeneity makes permutations of complete module lengths equivalent in aggregate risk, so each enumerated partition is realized as contiguous lengths without loss. No balanced-optimum assumption is used. Joint algorithm is the paper's original construction: for each module count, balance the lengths, assign margin $y_i = m_j(c + \varepsilon/\delta)$ inside module j, and select the least actual commitment. The exact joint oracle uses the already certified global optimum over structures and arbitrary bonuses. It is not substituted for the constructive algorithm.

Computation and certificates. For each scalar problem, monotone bisection uses the exact truncated recurrence, with 64 rational bisection steps after bracketing. At the retained lower margin the exact risk exceeds δ; the upper margin is feasible. We check 942 scalar partition problems (reused across prices), with maximum relative bonus-bracket width below 1.02×10^{-19}. We independently replay all 1,560 existing convex certificates using exact rational arithmetic. Restricting these exhaustive paid-prefix patterns to one module supplies the fixed optimized baseline, including possible unpaid suffixes. The global oracle intervals have relative width below 3.83×10^{-7}. Every reported risk is recomputed from the actual bonus vector and module boundaries. Ratios in the tables divide feasible implemented commitments by the corresponding feasible reference witness; machine-readable output also retains denominator-optimum intervals. All costs are stipulated game commitments, not tokens or API cash.

Empirical comparison. Table 6 compares the construction with the optimized alternatives. Its commitment exceeds the optimized structure–scalar baseline in all 240 instances, and exceeds fixed optimized rewards in 120. The construction's role is a certified approximation; exact optimization achieves lower commitment on these comparisons. The structure–scalar optimizer selects balanced partitions throughout this grid, as observed by exhaustive enumeration. The *No mechanism* baseline remains infeasible, with its literal zero commitment retained only for accounting; its normalized costs are undefined. Targets .8 and .9 have identical displayed cost aggregates because the inherited grid co-varies $\varepsilon = \alpha\delta$ and keeps ε/δ fixed. They are separate game settings, not a fixed-ε target-sensitivity experiment.

Reproducibility. The companion artifact retains the frozen protocol, source hashes, all scripts, and complete results in `controlled_ablation/`. The standard-library script `run_ablation.py` reproduces the computations; a separate checker replays exported risks and accounting, and a table script regenerates the manuscript inputs. The 1,440-row CSV retains target, worst risk, feasibility, module lengths, bonus vector, bonus/fee/total commitments, and both normalizations. Exact rational bonus witnesses and all scalar brackets are supplied. Target-specific feasible fractions and rounded aggregate ratios coincide with the pooled table (120 instances per target).

Table 5: Prespecified absolute examples: $n = 12$, $c = 1$, $\varepsilon = \delta$, $\rho/A = .04$ ($\rho = 69.12$). Bonus and fee are separated. An asterisk marks infeasibility. Rounding of feasible risk to its target does not indicate a violation; feasibility is checked exactly.

Target	Variant	Risk	Modules	Bonus	Fee	Total
0.8	No mechanism	1.00000*	1	0.00	0.00	0.00
0.8	Equal reward	0.20000	1	1543.40	0.00	1543.40
0.8	Reward-only (primary)	0.20000	1	1465.11	0.00	1465.11
0.8	Structure + equal reward	0.20000	4	514.00	207.36	721.36
0.8	Joint algorithm (ours)	0.08125	6	720.00	345.60	1065.60
0.8	Exact joint oracle	0.20000	4	503.73	207.36	711.09
0.9	No mechanism	1.00000*	1	0.00	0.00	0.00
0.9	Equal reward	0.10000	1	1543.40	0.00	1543.40
0.9	Reward-only (primary)	0.10000	1	1465.11	0.00	1465.11
0.9	Structure + equal reward	0.10000	4	514.00	207.36	721.36
0.9	Joint algorithm (ours)	0.04063	6	720.00	345.60	1065.60
0.9	Exact joint oracle	0.10000	4	503.73	207.36	711.09

Table 6: Controlled mechanism ablation on all 240 certified instances, at targets $1 - \delta \in \{.8, .9\}$. Commitment includes bonuses and restructuring fees. Ratios compare feasible commitments; parentheses give the worst ratio. The oracle is an exact-optimization reference, distinct from our constructive algorithm.

Variant	Feasible	C/C_{fixed} median (worst)	C/C_{oracle} median (worst)	Modules median
No mechanism	0/240	—	—	1
Equal reward	240/240	1.053 (1.077)	1.564 (6.104)	1
Reward-only (primary)	240/240	1.000 (1.000)	1.491 (5.667)	1
Structure + equal reward	240/240	0.689 (1.077)	1.018 (1.077)	2.5
Joint algorithm (ours)	240/240	0.990 (2.457)	1.500 (2.457)	4
Exact joint oracle	240/240	0.679 (1.000)	1.000 (1.000)	2.5

G HETEROGENEOUS PRIMARY BENCHMARK

Frozen primary design. The panel and protocol were locally frozen before its optimization runs; this is not an external preregistration. Its promotion to the main table in the current manuscript is an editorial decision made after observing the broader study. All other frozen panels and both positive and negative application diagnostics remain in the internal research archive. We froze all 300 instances before optimization: $n \in \{6, 8, 10, 12\}$ crossed with five cost patterns, five verification-weight patterns, and three tolerance patterns. Increasing and decreasing raw costs interpolate between $1/4$ and 4; alternating costs use these endpoints. The expensive and cheap bottleneck patterns place respectively 8 and $1/4$ at zero-based position $\lfloor n/2 \rfloor$, with raw cost one elsewhere. Each cost vector is normalized to mean one. Raw verification weights are uniform, front-heavy $(n, \ldots, 1)$, back-heavy $(1, \ldots, n)$, alternating $(1, 4, \ldots)$, or a middle critical item of weight eight among weights one; they are normalized to sum one. Raw tolerances are uniform, one middle noisy service of weight four among ones, or alternating $1/2, 2$; their mean is normalized to .01.

All instances use $\delta = .01$ and $\rho = n^3/25$. This stringent target is a computational design choice made before solving: $\delta < \min_i d_i$ in every instance, so a feasible contract cannot have a saturated item or an unpaid suffix. It permits exhaustive heterogeneous optimization without enumerating prefix truncations. These high-reliability games use one prescribed fee scaling; Appendix H separately varies targets and prices. Table 7 compares this grid with the historical homogeneous grid, which uses other targets and prices; the comparison is descriptive rather than a matched causal estimate of heterogeneity.

Distinct ablations. Fixed plus equal reward minimizes a single common bonus on the original chain; fixed plus optimized reward minimizes arbitrary bonuses on that chain. Structure plus reoptimized scalar minimizes a common bonus separately for each ordered partition and is the strong structure-only comparator. We also retain structure plus inherited scalar: it holds the optimal fixed-chain scalar bonus unchanged while choosing cuts. This last comparator is deliberately distinguished from the strong scalar comparator. Here it degenerates to the fixed chain, since that scalar already meets the target and extra cuts add positive fees without reducing the fixed bonus commitment. Joint optimization chooses both partition and arbitrary bonuses. The homogeneous constructive approximation theorem does not apply to the heterogeneous vectors; we report this method as inapplicable rather than extrapolating its guarantee.

Complete partition enumeration and convex formulation. We enumerate every one of the 2^{n-1} ordered contiguous cut masks. Unlike in the homogeneous reduction, permutations of module lengths are not identified. Because $\delta < \min_i d_i$, every feasible mechanism has $Q_i < 1$ and $y_i = b_i d_i - c_i > 0$. The exact untruncated recurrence is convex in these margins. For a fixed partition, the problem is therefore

$$\min_{y_i > 0} \sum_i \frac{c_i + y_i}{d_i} \quad \text{subject to} \quad F_R(y) := \sum_i d_i Q_i(y) \le \delta.$$

The risk constraint itself implies every terminal constraint $Q_i < 1$, so none is omitted by dropping separate terminal inequalities. Scalar bonuses are solved by monotone bisection, with both endpoints independently checked using rational arithmetic.

Rigorous pruning. For an item i, let W_i be the sum of verification weights from i through the end of its current module. The recurrence gives $F_R(y) \ge \sum_i \varepsilon_i W_i / y_i$. Weighted Cauchy–Schwarz therefore supplies the partition lower bound

$$\sum_i \frac{c_i}{d_i} + \frac{\left(\sum_i \sqrt{\varepsilon_i W_i / d_i}\right)^2}{\delta} + \rho(|R| - 1).$$

We round every square root downward to an exact rational multiple of 10^{-12} before evaluating this bound. A partition is pruned only if its lower bound is no smaller than a known feasible total commitment. Its bound and mask remain in the retained artifact output, allowing complete coverage to be checked independently.

Rational optimization certificates. SLSQP proposes candidate margins for the remaining partitions; its status is not treated as a certificate. Let y^+ be an exactly verified feasible point, $U = \sum_i y_i^+/d_i$, and $\lambda \geq 0$. Define

$$L(y;\lambda) = \sum_i y_i/d_i + \lambda(F_R(y) - \delta), \qquad h = \nabla_y L(\bar{y};\lambda).$$

An optimizer whose margin objective is at most U lies in $0 < y_i \leq d_i U$. Convexity then yields the certified margin lower bound

$$\max\left\{0,\ L(\bar{y};\lambda) + \sum_i \min\{-h_i\bar{y}_i,\ h_i(d_iU - \bar{y}_i)\}\right\}.$$

All recurrence values, derivatives, multipliers, feasibility comparisons, and tangent bounds are independently replayed as exact fractions. Adding $\sum_i c_i/d_i$ and the cut fee yields partition bounds; minima over all partition lower and upper bounds enclose the unrestricted joint optimum. We retain non-successful optimizer attempts rather than deleting their instances. The primary run contains 204,000 ordered partitions, 31,333 convex certificates and 172,667 rigorous prunes. All 204,000 scalar brackets pass independent rational replay. The largest joint relative interval width is 3.72×10^{-6}, below the frozen 10^{-5} tolerance. The 5,050 non-successful optimizer attempts do not invalidate these independently certified bounds.

Primary results. The strong structure-plus-scalar to joint ratio has median 1.221, minimum 1.001 and maximum 2.645 across the 300 instances. Fixed plus optimized rewards divided by joint has median 2.247, range $[1.283, 3.676]$. Joint commitment is more than 5%, 10% and 25% below the strong scalar comparator in respectively 232, 195 and 120 instances. These classifications use the less favorable ends of the optimum intervals, and no threshold comparison is unresolved. All 300 instances are included; these frequencies describe a deterministic grid, not population-level probabilities.

Post-hoc coupling diagnostic. After the primary grid, a separately frozen diagnostic holds each scalar optimizer's selected partition fixed and reoptimizes all individual bonuses. Its median commitment ratio to joint is 1.029 (maximum 1.435); the selected partition remains certified suboptimal for arbitrary rewards in 205/300 instances, with certified savings above 10% in 77/300. All 300 additional convex certificates pass rational replay (maximum relative interval width 5.59×10^{-6}). Thus reward reallocation on the scalar-selected structure recovers much of the 18.1% headline gap; a smaller residual structure–reward coupling remains. This diagnostic is post hoc, not an additional primary cohort. Separately frozen homogeneous matched controls and their complete results remain in the companion artifact.

Reproduction material. The `heterogeneous/` directory contains the primary protocol, frozen exact-rational parameters and hashes, all ordered-partition certificates, five implemented-mechanism rows per instance, scalar brackets, independent checks, figures and their source tables. Separately frozen post-hoc diagnostics reside in `posthoc_diagnostics/`. Exhaustive enumeration grows exponentially in n; the strict-target construction is specified above. All quantities are stipulated game commitments, distinct from model tokens or API prices.

H H2 HETEROGENEITY ROBUSTNESS

H2 uses the original cost, verification-weight, and regret-tolerance pattern definitions, with a deterministic reduced grid fixed before any new optimization. For every $n \in \{6, 8, 10, 12\}$, all five cost and three tolerance patterns are retained. The weight index is $(j_c + j_\varepsilon) \bmod |\mathcal{W}_n|$, where j_c, j_ε are zero-based original pattern indices. The eligible lists are $\mathcal{W}_6 = \{\text{uniform}, \text{alternating}, \text{critical}\}$ and $\mathcal{W}_n = \{\text{uniform}, \text{critical}\}$ for $n \in \{8, 10, 12\}$. This yields 60 base patterns. Eligibility is the strict, outcome-independent condition $\min_i d_i > .05$; front/back-heavy weights are omitted, and the alternating $n = 8$ pattern is excluded at equality. The numerical framework thus remains within the unsaturated regime rather than silently extrapolating its convex program to saturated optima.

Each vector is interpolated componentwise as $x(\lambda) = (1-\lambda)x_{\text{hom}} + \lambda x_{\text{het}}$, simultaneously for c, d, ε, with $\lambda \in \{0, 1/4, 1/2, 3/4, 1\}$ and homogeneous values $c_i = 1$, $d_i = 1/n$, $\varepsilon_i = .01$.

Table 7: Commitment divided by joint optimized commitment: median (worst) across every grid case. The homogeneous grid (240 grid cases) and heterogeneous grid (300 prespecified grid cases) use different frozen parameter regimes; their difference is descriptive, not a matched causal estimate of heterogeneity. Arbitrary-bonus optima have independently certified numerical intervals.

Method	Homogeneous	Heterogeneous
Fixed + equal	1.564 (6.104)	2.570 (4.755)
Fixed + optimized	1.491 (5.667)	2.247 (3.676)
Structure + optimized scalar	1.018 (1.077)	1.221 (2.645)
Constructive theorem algorithm	1.500 (2.457)	Not applicable
Joint optimized reference	1.000 (1.000)	1.000 (1.000)

Exact normalization is preserved without further adjustment. This is a matched joint continuation of the three vectors, not separate attribution to cost, weight, or tolerance heterogeneity. Risk targets are $\delta \in \{.01, .025, .05\}$ and per-cut fees are $\rho \in \{n^3/100, n^3/25, n^3/5\}$. Since $\min_i d_i(\lambda) \geq (1-\lambda)/n + \lambda \min_i d_i(1) > .05$, every case meets the strict unsaturation condition. The full 2,700-case index contains 2,196 unique scientific inputs; repeated homogeneous endpoints are solved once and reused by exact input identity.

For every ordered partition, the scalar comparator reoptimizes its common bonus, while the joint method permits arbitrary individual bonuses. The fixed-organization comparator permits arbitrary rewards on the uncut chain. All ratios retain outward-enclosing optimum intervals. A certified $> 10\%$ improvement means $1 - U_{\text{joint}}/L_{\text{scalar}} > .1$, not merely a ratio exceeding 1.1. Different selected cut masks are descriptive witnesses; the stronger statistic requires the lower bound for the scalar-selected mask after arbitrary reward optimization to exceed the global joint upper bound. It does not establish uniqueness or exclude tied scalar-optimal masks. Table 2 summarizes the continuation, and Figure 6 shows its target- and fee-dependent results.

Independent rational replay verified 1,493,280 ordered-partition bounds, 362,680 convex certificates and 1,493,280 scalar brackets across all 2,196 unique inputs. The largest global relative interval width was 4.64×10^{-6}; no comparison uses an optimizer success flag as proof.

The protocol, initial and amended solver versions, and all non-success attempts remain in the internal research archive. Table 8 reports the endpoint sensitivity for this restricted-weight grid; all nine evaluated target/fee settings show increasing median ratios along the five strengths.

Table 8: H2 endpoint sensitivity. Each endpoint cell contains 60 base patterns; all nine median-ratio sequences increase monotonically across the five strengths.

δ	ρ	Median ratio, $\lambda = 0$	Median ratio, $\lambda = 1$	$> 10\%$, $\lambda = 1$
.01	$n^3/100$	1.000	1.147	35/60
.01	$n^3/25$	1.008	1.104	27/60
.01	$n^3/5$	1.022	1.075	24/60
.025	$n^3/100$	1.002	1.271	42/60
.025	$n^3/25$	1.008	1.159	35/60
.025	$n^3/5$	1.011	1.121	30/60
.05	$n^3/100$	1.002	1.359	43/60
.05	$n^3/25$	1.009	1.205	39/60
.05	$n^3/5$	1.023	1.167	33/60

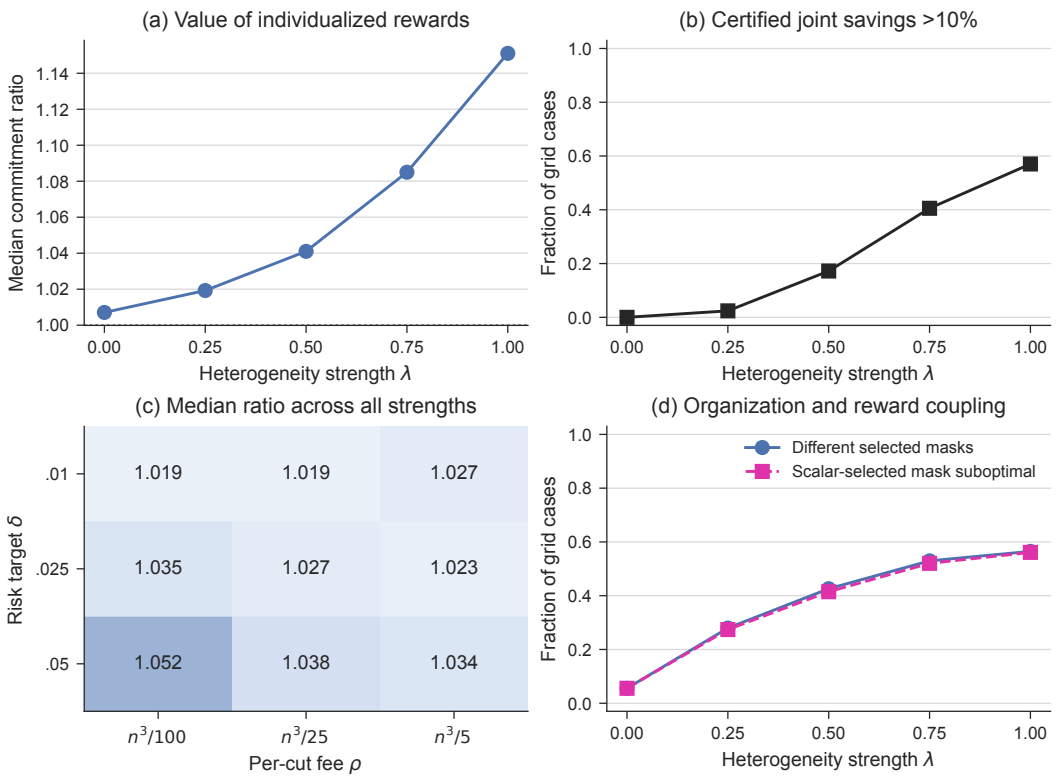

Figure 6: Frozen H2 robustness map. Panels (a,b,d) pool the nine target/fee settings within each strength; panel (c) pools all five strengths. Certified mask loss checks whether the particular scalar-selected structure remains suboptimal even after arbitrary reward reoptimization; tied scalar-optimal alternatives are not excluded. Exact per-strength, per-target, and per-fee CSV summaries accompany the package.

I E6 PROTOCOL AND COMPLETE TRAJECTORIES

The behavioral bridge instantiates the one-shot game exactly, then uses repeated choices to form an empirical distribution. It does not assume a repeated-game equilibrium. Actions are $a_i \in \{0, 1\}$, with $L = 0, H = 1$ and no decline action. The two structures have success probabilities

$$p_D(a) = \tfrac{1}{2}a_1 + \tfrac{1}{2}a_1a_2, \qquad p_B(a) = \tfrac{1}{2}a_1 + \tfrac{1}{2}a_2, \tag{32}$$

and expected stage utility $u_i(a) = b_i p_R(a) - a_i$. At each round an independent uniform draw selects one of the two local verification items; only that item's success determines the shared success event and each agent's own bonus. The effort cost is paid regardless of success. The resulting expected utility is exactly u_i. Bonuses and costs are proxy credits, distinct from API expenditure. The principal's cut fee affects commitment, not agents' stage utilities. There is no additive participation payment or monetary transfer to a model.

Frozen design and optimizer certificate. We choose $\delta^* = .10$, $\varepsilon^* = .05$ and $\rho = .25$ before any calls, solely from the exact finite game's geometry. The fixed-chain optimum is $11/2 + \sqrt{6}$, attained at bonuses $3 + \sqrt{6}/2$ and $5/2 + \sqrt{6}/2$. We round both outward to twelve decimal places: $(4.224744871392, 3.724744871392)$. The commitment is 7.949489742784 and certified risk $.09999999999997376$. The joint optimum cuts the edge, uses $(3, 3)$, and costs $6 + .25 = 6.25$, with risk $.10$. The under-incentivized chain uses $(.5, .5)$, commitment one, and worst risk one. Since $\delta^* < \min_i d_i$, every feasible optimized prefix is unsaturated. Thus enumerating both structures and optimizing their full bonus vectors does not omit a cheaper saturated or zero-payment solution. An algebraic solution and rational convex tangent bounds certify the fixed and joint optima, with a certified upper–lower gap of about 1.01×10^{-12}. Independent four-profile linear programs check the recurrence, including its zero-tolerance boundary. The design tests the exact theorem and optimizers, not the homogeneous scaling law, whose assumption $\varepsilon \geq c\delta$ is not satisfied here.

Cohorts and information. The homogeneous pairs are DeepSeek-V4.1-Flash (DeepSeek-AI, 2026) (served identifier `deepseek-flash`) and Qwen3.6-35B-A3B (Qwen Team, 2026) (`qwen3.6-35b-a3b`). Each condition has four prespecified fresh-conversation trajectories, each with 64 rounds. This gives 24 planned trajectories and 3,072 decision slots. No mixed pair or extension to 128 rounds is used. Distinct frozen seeds control local verification and caller order, not provider sampling. Each agent has its own persistent logical conversation, implemented by resending its message history. Prompts disclose public production rules, own bonus and cost, and own realized feedback. They contain no payoff table, expected or counterfactual utility, regret, best-response label, equilibrium hint, opponent response or direct opponent action. Both actions commit before either agent receives feedback, even though transport calls occur sequentially in randomized order. Outcome feedback may naturally reveal information about effort through the production rules.

Feedback contains the agent's dispatched action, sampled item, success, own bonus, effort cost, net and cumulative payoff. Decimal strings preserve arithmetic. Final-round feedback is archived without requesting a 65th action. Temperature is .2, the response limit is 128 tokens, and thinking is disabled. The parser accepts only an object with the single key `effort` equal to `L` or `H`. Malformed, extra-key, non-stop, reasoning-bearing or over-cap output maps to L without retry; actual model text remains in history. No such invalid response occurred in the 3,072 successful calls retained here. Exact prompts, request payloads, raw streaming responses and parsed outputs are retained in the internal research archive.

Prespecified regret and reliability. For a trajectory prefix of length T, define

$$\widehat{r}_i(T) = \max_{x \in \{0,1\}} \frac{1}{T} \sum_{t=1}^{T} [u_i(x, a^t_{-i}) - u_i(a^t)], \tag{33}$$

$$\widehat{\varepsilon}_i(T) = \max\{0, \widehat{r}_i(T)\}, \qquad \widehat{\varepsilon}_{\max}(T) = \max_i \widehat{\varepsilon}_i(T). \tag{34}$$

Signed raw regrets are retained. The nonnegative part supplies an admissible theorem tolerance. The evaluator uses expected stage utilities, not sampled counterfactual rewards, and never supplies these calculations to the agents. At checkpoints $8, 16, 32, 64$, entry means $\widehat{\varepsilon}_{\max}(T) \leq .05$; reliability success means $T^{-1} \sum_t p_R(a^t) \geq .90$. We separately retain realized sampled failure, action

frequencies, first entry, and persistence at subsequent observed checkpoints. Persistence does not assert that every intervening round is within tolerance. The empirical risk is necessarily bounded by the exact risk at its measured vector tolerance, and also by the scalar envelope. The scientifically meaningful observation is entry into the tolerance fixed before play, jointly with reliability under lower commitment.

Observed outcomes. Table 9 reports all 24 trajectories. DeepSeek-V4.1-Flash completes all twelve trajectories. Its eight optimized trajectories have only HH profiles, zero regret and local reliability one at every checkpoint. Control final regrets are $0, .4921875, .01171875, .03125$, and reliability lies between zero and $.03125$. Qwen3.6-35B-A3B completes all twelve trajectories: 4/4 fixed and 4/4 joint trajectories meet both targets, while 0/4 controls enter the tolerance and 0/4 meet reliability. All eight control trajectories fail the reliability target. The optimized chain gives service 1 a strict high-effort incentive; service 2 still depends on upstream effort. Both cut-structure services have strict high-effort incentives. Across fresh trajectories, both model families consistently choose the high-effort action under the optimized contracts. The contrast changes both optimized bonuses and structure, not a factorial decomposition.

Table 9: All 24 E6 trajectories. A = optimized fixed chain; B = optimized joint cut; C = under-incentivized chain. First entry uses the prespecified checkpoints; – means no observed entry. Displayed values are rounded; full precision is retained in the artifact.

Model	Condition/replicate	Rounds	First entry	Final regret	Final reliability
DeepSeek-V4.1-Flash	A/0	64	8	0.0000	1.0000
DeepSeek-V4.1-Flash	A/1	64	8	0.0000	1.0000
DeepSeek-V4.1-Flash	A/2	64	8	0.0000	1.0000
DeepSeek-V4.1-Flash	A/3	64	8	0.0000	1.0000
DeepSeek-V4.1-Flash	B/0	64	8	0.0000	1.0000
DeepSeek-V4.1-Flash	B/1	64	8	0.0000	1.0000
DeepSeek-V4.1-Flash	B/2	64	8	0.0000	1.0000
DeepSeek-V4.1-Flash	B/3	64	8	0.0000	1.0000
DeepSeek-V4.1-Flash	C/0	64	–	0.4922	0.0312
DeepSeek-V4.1-Flash	C/1	64	64	0.0312	0.0000
DeepSeek-V4.1-Flash	C/2	64	8	0.0000	0.0000
DeepSeek-V4.1-Flash	C/3	64	16	0.0117	0.0156
Qwen3.6-35B-A3B	A/0	64	8	0.0000	1.0000
Qwen3.6-35B-A3B	A/1	64	8	0.0000	1.0000
Qwen3.6-35B-A3B	A/2	64	8	0.0000	1.0000
Qwen3.6-35B-A3B	A/3	64	8	0.0000	1.0000
Qwen3.6-35B-A3B	B/0	64	8	0.0000	1.0000
Qwen3.6-35B-A3B	B/1	64	8	0.0000	1.0000
Qwen3.6-35B-A3B	B/2	64	8	0.0000	1.0000
Qwen3.6-35B-A3B	B/3	64	8	0.0000	1.0000
Qwen3.6-35B-A3B	C/0	64	–	0.3750	0.5078
Qwen3.6-35B-A3B	C/1	64	–	0.3789	0.2734
Qwen3.6-35B-A3B	C/2	64	–	0.4922	0.0312
Qwen3.6-35B-A3B	C/3	64	–	0.6211	0.7578

Experimental provenance. All 24 trajectories and 3,072 actions are retained. Five unsuccessful HTTP 429 attempts and the transport-only pacing/backoff amendment are indexed in the artifact; every reused successful response had an identical request history. Independent raw-history reconstruction and rational finite-game enumeration reproduce all 96 checkpoints. Requested and served identifiers are provider aliases, not immutable weight checkpoints; provider generation is not deterministic. Exact prompts, payloads, raw responses, parser outputs, request accounting and latency measurements are retained in the internal research archive.

J REPRODUCIBILITY AND ARTIFACT INVENTORY

The manuscript source bundle contains all rendered figures, plotted data, result-figure generators, source hashes and compilation instructions. Figures 1 and 2 are supplied PNG schematics; Figures 3 and 4 retain editable vector sources. The minimal code supplement, also available at `https://anonymous.4open.science/r/anonymousgit-BBF2/`, contains the scientific mechanism, frozen input generators and compact main-comparison records.

Material	Included reproduction support
Mechanism and optimization	Exact-risk recurrence, homogeneous construction, ordered-mask optimization, independent rational certificate checker.
300 primary instances	Exact inputs and bounds, feasible contracts, all optimized comparators and post-hoc reward-reallocation diagnostics.
H2 continuation	All 2,196 distinct input records; generator for the 2,700-entry index; five-strength replay (target/fee summary tables are in the manuscript source bundle).
240 homogeneous checks	All six comparators, 1,440 mechanism records and exact feasible bonuses.
E6 behavioral game	All 24 action trajectories and 96 checkpoints; frozen verification and call-order streams; scoring, prompts and opt-in live runner.
Representative certificates	Two complete heterogeneous examples and one homogeneous paid-prefix example, plus per-instance regeneration commands.

From the extracted code directory, `python -m structpay.reproduce` recomputes the paper summaries without network access using the Python standard library. It checks record hashes, feasible contracts, E6 regret and reliability, and the representative rational certificates. Replay uses cached intervals for other comparisons; it does not independently re-establish every cached lower bound. Fresh optimization with NumPy/SciPy regenerates full per-instance certificates, starting with one case:

```
python -m structpay.benchmarks --suite primary --limit 1
python -m structpay.benchmarks --suite h2 --limit 1
python -m structpay.benchmarks --suite homogeneous --limit 1
```

The packaged heterogeneous solver covers positive tolerances and $0 < \delta < \min_i d_i$, which includes the primary and H2 panels. The exact-risk recurrence has the broader domain of Theorem 1. Complete-panel limits and dependency versions are given in the code README. Optional live E6 runs use environment-supplied credentials; fresh provider generations remain stochastic.

The internal research archive retains exhaustive optimizer certificates, original provider payloads, accounting and latency records, full historical protocols, and additional behavioral and application diagnostics with both positive and negative outcomes. These records are distinct from the compact code supplement. The manuscript's hash-indexed evidence inventory identifies 347 retained source files; it is a provenance index, not a public repository link or a claim that all archived files are embedded in this package.