TOKEN AS POWER: HOW TOKEN RESOURCES SHAPE AGENT ECONOMIES

Anonymous authors
Paper under double-blind review

ABSTRACT

In evolving multi-agent economies, task performance governs economic rewards, evolutionary survival, and reproduction. However, whether intrinsic model capability guarantees competitive dominance under asymmetric resource constraints remains an open question. Motivated by debate experiments where a compute-advantaged weaker model defending an incorrect claim wins 83% of debates at matched monetary cost, we systematically investigate token allocation within the Economy of Minds (EOM) framework across two scientific-reasoning benchmarks. We examine three regime types: persistent resource inequality, temporary budget shocks, and wealth-funded token expansion. Our investigation reveals three core dynamics. First, resource-advantaged weaker models can displace models that perform better under matched budgets, a phenomenon we term the *Agentic Gresham-like Effect*, even when stronger models retain a static capability advantage. Second, temporary budget shocks cause amplified losses under economic evolution compared to static populations, leaving performance gaps and changes in population and wealth after budget restoration. Third, early token investment can support subsequent lineage expansion: restricting purchases by selected agents reduces their lineages' final population shares in paired comparisons. Together, these findings establish token access as a structural driver of population dynamics rather than a passive runtime hyperparameter. Consequently, evaluating agent economies requires reporting resource distributions alongside task performance and lineage trajectories.

1 INTRODUCTION

Large language model (LLM) agents can solve complex tasks through debate and collaboration (Du et al., 2023; Wu et al., 2023; Hong et al., 2024). Recent work extends agent interactions to simulated markets and auction-based service allocation (Zhao et al., 2024; Yang et al., 2025). **Crucially, evolving agent economies can link current task performance to subsequent survival and reproduction.** In Economy of Minds (EOM), agents receive economic rewards for task contributions, while accumulated wealth affects their ability to pay operational rents and their opportunities to reproduce; bankrupt agents are removed (Qi et al., 2026). Task performance can therefore influence not only current rewards but also which lineages remain represented in future interactions.

However, does stronger model capability necessarily confer a competitive advantage that supports lineage survival? Our preliminary debate experiment on StrategyQA Geva et al. (2021) illustrates how token budgets can reverse competitive outcomes. DeepSeek-V4-Pro DeepSeek-AI et al. (2026) defended the correct answer with an 800-token budget, while DeepSeek-V4-Flash defended the incorrect answer with budgets ranging from 800 to 3,200 tokens. GPT-4o Yenduri et al. (2023) judged 100 debates per condition. At equal token budgets, Flash won 44% of debates; its win rate rose to 83% at 3,200 tokens, matching the monetary cost of Pro's 800-token budget (Figure 1).

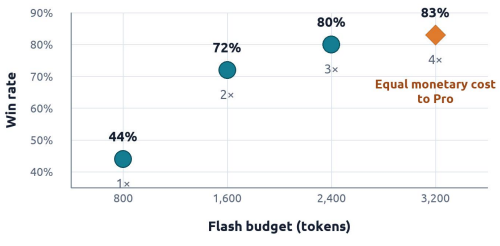

Figure 1: StrategyQA debates: Flash defended the incorrect answer against Pro (800 tokens). GPT-4o judged 100 debates per condition. †Equal monetary cost to Pro.

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.

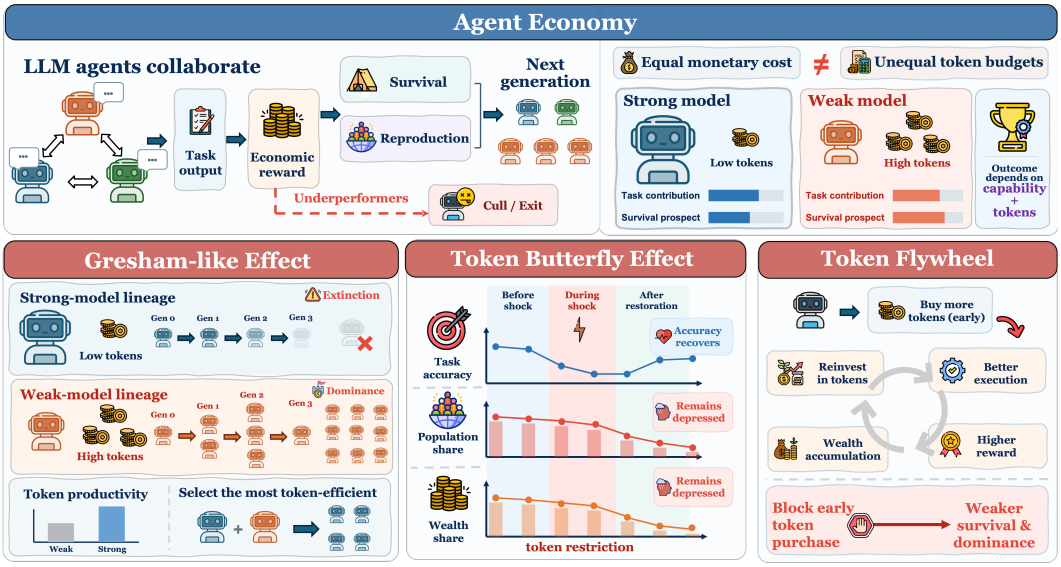

Figure 2: Overview of token resources in an evolving agent economy. Top: illustration of how token resources enter the economic feedback between task contributions and reproduction. Bottom: (1) Unequal budgets can reverse competitive outcomes. (2) Temporary restrictions can leave effects after budgets are restored. (3) Wealth-funded token investment can support subsequent lineage expansion.

Thus, under equal monetary costs, the weaker model won more often despite defending the incorrect answer. This reversal motivates our investigation of whether token advantages also translate into lineage survival and expansion in an evolving agent economy.

This phenomenon brings to mind Gresham's Law. In classical economics, Gresham's Law states that when two forms of money have equal legal tender value but different intrinsic values, people tend to spend the lower-value money and hoard the higher-value money, driving the latter out of circulation. A clear similarity emerges in our evolutionary market: high-budget lineages of weaker models survive and proliferate, while low-budget lineages of stronger models exit the market. We term this phenomenon an *Agentic Gresham-like Effect*. The analogy lies strictly in the displacement phenomenon itself, rather than in identical monetary or legal tender mechanisms.

Prior research on multi-agent debate and collaboration focuses primarily on improving task performance through agent interaction (Du et al., 2023; Wu et al., 2023; Hong et al., 2024). The Gresham-like effect motivates us to examine how token resource differences shape competition and lineage survival under economic feedback. We therefore ask:

Core Question

Can externally assigned resource inequality change the outcome of economic competition? If these externally imposed inequalities are only temporary, do they still leave persistent historical marks? Furthermore, if resources are no longer passively assigned but can be actively acquired through agent investment using earned profits, does this mechanism generate sustained feedback loops?

We design three major categories of dynamic agent economy experiments to examine how token variations affect lineage outcomes. First, within the Economy of Minds (EOM) framework, we conduct static cost-equivalent evaluations between strong and weak models, subsequently tracking the long-term impact of token allocation on model competition under dynamic evolutionary settings. Second, we introduce budget restrictions in the agent market to analyze the downstream impacts of transient resource shocks, followed by recovery experiments to test whether restoring token budgets allows market structures to self-repair. Third, moving beyond passive resource limits, we implement an active investment mechanism where agents can autonomously choose to purchase additional tokens prior to taking action, allowing them to recognize the strategic value of compute and make

investments that shape future market outcomes. Our code is available at [1] and will be publicly available when published.

Our empirical results uncover three primary dynamics:

- **Agentic Gresham-like Effect** (Section 4): Both in static tasks and dynamic evolutionary environments, weaker models equipped with higher token budgets survive, reproduce, and establish dominant lineages and offspring, while overall team task accuracy remains largely stable. To relieve this, we introduce **Token Productivity (TP)**, which selects parents using contribution quality normalized by token consumption rather than accumulated wealth. TP yields higher mean performance across evaluation checkpoints in all four reported settings.
- **Token Butterfly Effect** (Section 5): Under equivalent token restrictions, the performance gap is larger in an active agent market than in a fixed population baseline: $B1 - B2 > A1 - A2$ during the reduction window. Restoring token budgets can leave gaps in task performance and in the population and wealth of affected roles, producing persistent consequences beyond the intervention period.
- **Token Flywheel** (Section 6): Paired interventions reveal that early token purchases can alter lineage survival by triggering a self-reinforcing feedback loop. Weighted rewards drive wealth accumulation and compute reinvestment, compounding transient advantages into persistent dominance for Agent lineages. However, instances of token-purchasing agents still exiting show that compute investment alone is not a sufficient guarantee of success.

These findings demonstrate that token allocation must be conceptualized as a core evolutionary mechanism within agent economies rather than a mere computational hyperparameter. Because resource disparities shape immediate outputs and dictate long-term lineage survival through economic feedback, market dominance cannot be conflated with genuine collective capability or long-term adaptability under shifting task demands. Future mechanism design must therefore jointly optimize resource allocation alongside reward and reproduction rules, not to artificially shield strong models or preserve every lineage, but to foster agent economies that excel at immediate tasks while retaining the latent capabilities needed to dynamically adapt as operational requirements evolve.

2 Related Work

Multi-Agent Collaboration and Debate LLM agents improve task performance through debate, role specialization, and output aggregation (Du et al., 2023; Hong et al., 2024; Wang et al., 2024). Adaptive approaches optimize team composition and coordination (Liu et al., 2024; Dang et al., 2025), while scaling studies examine the effects of agent count and communication costs (Li et al., 2024a;c; Kim et al., 2026). We examine how resource differences affect lineage survival across tasks, rather than only the performance of a given team.

Agent Economies and Population Dynamics Economic agent systems study market competition, service allocation, and collective behavior (Zhao et al., 2024; Yang et al., 2025; Li et al., 2024b). EOM connects task rewards to wealth, survival, and reproduction (Qi et al., 2026), providing the setting for our experiments. Unlike evolutionary approaches that search for better agent designs (Hu et al., 2025; Zhang et al., 2026; Weng et al., 2026), we investigate how token inequality, temporary shocks, and investment shape selection within an economy.

Compute Allocation and Resource-Constrained Agents Prior work balances performance and cost through model routing, adaptive inference budgets, and communication pruning (Ong et al., 2025; Snell et al., 2024; Han et al., 2025; Zhang et al., 2024). Studies of overthinking and cost-aware evaluation also distinguish additional computation from useful computation (Chen et al., 2025; Kapoor et al., 2024). We extend this perspective to economic feedback: current performance affects wealth and reproduction, while token purchases can change future resource access. Our outcomes therefore include lineage dynamics alongside task performance.

Additional related work is discussed in Appendix B.

[1] https://anonymous.4open.science/r/token-as-power-6EBC/README.md

3 AGENT ECONOMY AND EXPERIMENTAL DESIGN

3.1 ECONOMIC ENVIRONMENT AND TOKEN RESOURCES

We use the scientific reasoning setting of Economy of Minds (EOM) (Qi et al., 2026), with Literature, Planner, Deriver, Verifier, and Answer agents.We adjust Deriver token budgets to intervene in a key reasoning stage. Each agent uses a model and a prompt that guides its actions. A founder and its descendants form a lineage. Several lineages can use the same model.

Agents bid for opportunities to contribute to a shared task context. After each training task, rewards update their wealth. Rent is charged at regular intervals, and agents with negative wealth leave the population. The system can create new agents after a bankruptcy or at regular intervals. By default, new prompts are made by changing the wealthiest agent's prompt or revising a failed prompt using task feedback and reference answers. New agents keep the role and lineage of their source agent. If a required role disappears, the system creates a replacement.

Token budgets determine how much agents can generate. Only purchasing experiments allow agents to spend wealth on larger budgets. Wealth is measured in internal market units and is separate from the monetary API cost used in the preliminary experiment.

3.2 EXPERIMENTAL DESIGN

We study persistent token inequality, temporary reductions followed by restoration, and wealth-funded token purchases. Each results section specifies its interventions. The StrategyQA preliminary experiment (Figure 1) compares debate win rates under matched token budgets and monetary costs. GPT-4o judges agents defending correct and incorrect answers; winning does not imply answer correctness.

Economic experiments use FrontierScience-Research (Wang et al., 2026) and a filtered text-only PHYSICS subset (Feng et al., 2025), with fixed evaluation sets of 19 and 20 tasks, respectively. PHYSICS additionally contains 30 non-overlapping evolution tasks. Filtering and dataset details appear in Appendix A.

Persistent-inequality experiments compare DeepSeek-V4-Flash/Pro and GPT-5.1/5.2 within the Deriver role. Other roles use Flash and GPT-5.1, respectively. "Stronger" and "weaker" refer to static performance under matched token budgets on the evaluated tasks. Shock and purchasing runs use one shared backbone across interacting roles; configurations and inheritance rules are specified in their respective sections.

3.3 EVALUATION PROTOCOL AND OUTCOMES

We evaluate task performance separately from ecological outcomes. Regular dynamic runs process 30 evolution tasks. For runs with held-out evaluation, we evaluate after every three tasks, producing ten checkpoints. At each checkpoint, the current population and economic state are frozen during evaluation on the fixed test set. Evaluation results do not update wealth or affect elimination and reproduction.

On FrontierScience, a response is counted as correct when its normalized rubric score is at least 0.7. On PHYSICS, we report the mean per-question score without additional thresholding. Let $s_{cj} \in [0,1]$ and $p_{cj} \in [0,1]$ denote the respective scores for question j at checkpoint c. We compute

$$\mathrm{Acc}_{\mathrm{FS}}(c) = \frac{100}{19}\sum_{j=1}^{19}\mathbb{I}[s_{cj} \geq 0.7], \qquad \mathrm{Score}_{\mathrm{PHYSICS}}(c) = \frac{100}{20}\sum_{j=1}^{20} p_{cj}. \tag{1}$$

Missing final answers receive zero and remain in the evaluation denominator. Scientific-reasoning experiments use GPT-5.2 for LLM-based judgments; PHYSICS answer evaluation combines symbolic equivalence checking with GPT-5.2 as a fallback. For runs with ten checkpoints, we additionally report the arithmetic mean of the checkpoint performance values.

We track population size, wealth, public action counts, and offspring counts to characterize competition within the economy. We report population and wealth shares for the roles, model groups, or founder lineages compared in each experiment.

4 Agentic Gresham-like Effect

We examine whether model performance under equal token budgets predicts survival under unequal resources, then test a token-productivity-based alternative to wealth-based reproductive selection.

Table 1: Static Deriver model–budget comparisons. FS denotes FrontierScience; metrics follow Section 3.

Model	Budget	FS acc. (%)	FS score	PHYSICS (%)
DeepSeek-V4-Flash	8,192	10.5	0.379	46.46
DeepSeek-V4-Pro	2,048	10.5	0.302	44.56
DeepSeek-V4-Pro	8,192	15.8	0.405	51.26
GPT-5.1	8,192	10.5	0.394	39.90
GPT-5.2	2,048	15.8	0.420	36.40
GPT-5.2	8,192	21.1	0.458	51.70

4.1 Static Capability Calibration

At the same 8,192-token budget, Pro outperforms Flash and GPT-5.2 outperforms GPT-5.1 on both benchmarks, establishing the stronger and weaker model groups. We then compare stronger models at 2,048 tokens with weaker models at 8,192 tokens, which incur comparable monetary costs under the API pricing used in our experiments. Under this allocation, the two groups achieve broadly comparable static performance, with their relative ranking varying across model–dataset settings (Table 1).

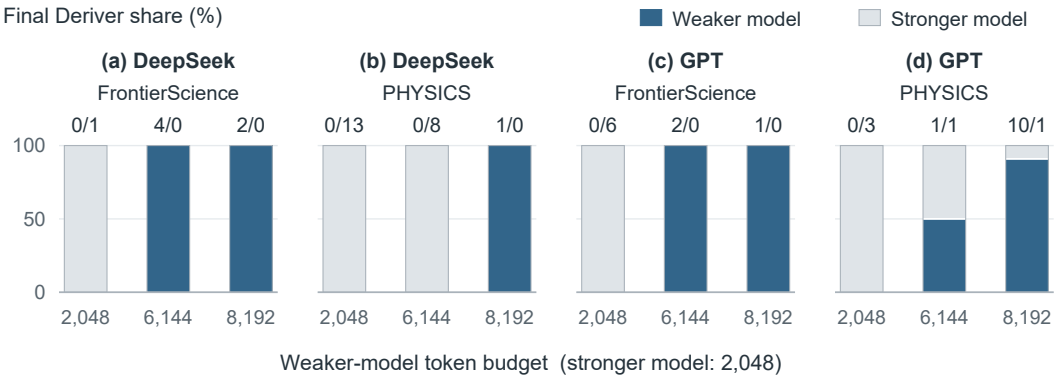

Figure 3: Final Deriver shares as the weaker model's budget grows; the stronger model remains at 2,048 tokens. Labels give weaker/stronger counts: Flash/Pro for DeepSeek and GPT-5.1/GPT-5.2 for GPT, aggregated across founder lineages.

4.2 Resource-Dependent Competitive Outcomes

Each market starts with two Derivers from each model group and one agent for each of the other four roles: Literature, Planner, Verifier, and Answer. Parent selection is based on wealth. The stronger model receives 2,048 tokens, while the weaker model receives 2,048, 6,144, or 8,192 in separate conditions. Descendants inherit their parent's model and token allowance.

At equal token budgets, only stronger-model Derivers survive in all four settings. As the weaker model's budget increases across the tested levels to 8,192 tokens, this pattern reverses: weaker-model Derivers completely displace their stronger counterparts in three settings and occupy ten of eleven final Deriver positions on GPT–PHYSICS (Figure 3). Thus, broadly comparable static performance at comparable monetary cost gives way to weaker-model dominance in dynamic competition. We call this resource-driven displacement the *Agentic Gresham-like Effect*.

These experiments show that token allocation shapes competitive outcomes and can reverse which model group dominates the surviving population. Detailed wealth, action, and held-out performance trajectories appear in Appendix D.

4.3 New selection method-Token Productivity

In order to relieve the *Agentic Gresham-like Effect* , we introduce *Token Productivity* (TP) to select Deriver parents based on contribution quality and resource use, replacing wealth as the selection criterion. TP normalizes contribution quality above a baseline by actual token consumption. Using at most three recent contributions, let $\widetilde{q}_i(t)$ denote mean contribution quality smoothed with one prior observation of score 0.5, and $\widetilde{u}_i(t)$ denote mean actual generation-token consumption. We define

$$\mathrm{TP}_i(t) = \frac{[\widetilde{q}_i(t) - 0.5]_+}{\max(1, \widetilde{u}_i(t)/2048)}, \qquad [x]_+ = \max(x, 0). \tag{2}$$

Mean token consumption affects the denominator only above 2,048 tokens; at or below that level, equal quality receives equal TP regardless of consumption. Agents without contribution history receive zero. Estimator definitions and selection details appear in Appendix E.

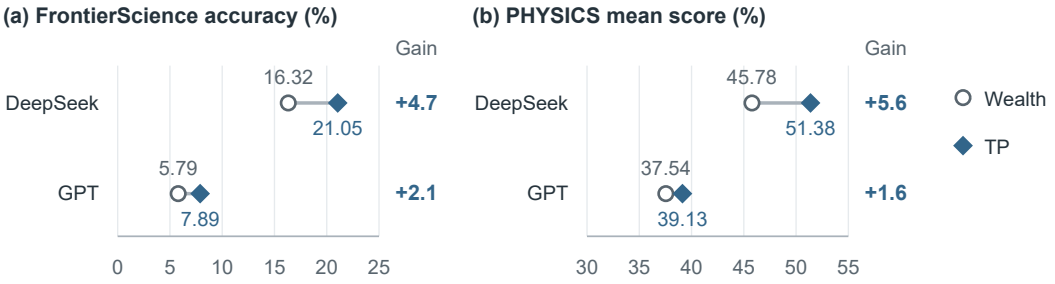

Figure 4: Ten-checkpoint mean performance under Wealth and TP selection. Panels use different scales for FrontierScience accuracy and PHYSICS mean score. Gains are TP minus Wealth in percentage points, computed before rounding.

TP yields higher mean checkpoint performance in all four reported comparisons (Figure 4): FrontierScience accuracy increases by 2.1–4.7 percentage points and PHYSICS mean score by 1.6–5.6. TP prioritizes agents with higher token productivity for reproduction, promoting population compositions that balance contribution quality and token consumption.

Table 2: Token-shock conditions. Fixed-population conditions disable economic feedback into agent selection; evolving conditions retain wealth updates, bankruptcy, and reproduction.

Shock Type	No	Temporary
Fixed population	A1	A2
Evolving economy	B1	B2

5 Token Butterfly Effect

We ask whether temporary token reductions have consequences beyond the restriction window. The *Token Butterfly Effect* refers here to observed amplification and subsequent effects.

5.1 Token Shock Design

We use a 2×2 design crossing economic evolution with a temporary token reduction (Table 2). All conditions branch from the same population and economic state at $T10$. A1/B1 retain 8,192-token allowances throughout. In A2/B2, Deriver allowances fall to 1,024 during $T11$–$T20$ and return to 8,192 at $T21$. Other roles remain at 8,192. Frozen evaluations use the allowance active at each checkpoint.

5.2 Performance Loss and Recovery

Let $\overline{P}_X$ denote condition X's mean performance over shock checkpoints $T12, T15, T18$. We compare losses under fixed populations, $\Delta_A = \overline{P}_{A1} - \overline{P}_{A2}$, and evolving markets, $\Delta_B = \overline{P}_{B1} - \overline{P}_{B2}$:

$$\Delta_{\mathrm{amp}} = \Delta_B - \Delta_A. \tag{3}$$

All four performance comparisons have positive contrasts (Table 3), consistent with amplification in the reported runs. These contrasts reflect how the dynamically evolving system amplifies the effects of token shock.

Table 4: Ecological states before the shock ($T10$), at its end ($T20$), and after recovery ($T30$). Population entries are Derivers/all living agents; wealth entries are Deriver wealth/total market wealth, rounded to two decimals.

	$T10$		$T20$		$T30$	
Condition	Population	Wealth	Population	Wealth	Population	Wealth
DeepSeek-V4-Flash / FrontierScience						
B1	10/16	3.58/5.29	4/14	1.36/4.66	5/20	1.82/8.48
B2	10/16	3.58/5.29	6/18	2.26/6.65	2/20	0.65/8.29
DeepSeek-V4-Flash / PHYSICS						
B1	2/12	0.87/6.23	1/16	0.85/8.23	1/20	0.33/6.94
B2	2/12	0.87/6.23	1/13	0.46/7.21	2/19	0.66/6.10
GPT-5.4 / FrontierScience						
B1	5/14	1.10/4.57	6/17	2.25/7.19	7/19	2.44/6.75
B2	5/14	1.10/4.57	7/16	3.07/6.71	3/19	0.58/6.57
GPT-5.2 / PHYSICS						
B1	7/14	3.40/5.55	4/17	1.30/6.87	3/20	0.65/5.00
B2	7/14	3.40/5.55	6/18	2.51/7.88	1/19	0.23/6.40

Also, during the recovery period, B2's mean performance remains below that of the unchanged-budget control B1 in all four evolving-market settings, leaving a persistent gap in system-level accuracy.

Beyond accuracy, delayed ecological scars can also emerge in population and wealth. On DeepSeek–FrontierScience, B2 has more Derivers than B1 at $T20$ (six versus four), but fewer at $T30$ (two versus five); its wealth ordering also reverses. Both GPT settings show a similar reversal in Deriver counts. See details in Appendix F due to page limits.

6 Token Flywheel Effect

Table 3: Shock-window losses and recovery-window gaps in percentage points. FS denotes FrontierScience accuracy; PHYSICS uses mean score. Recovery gaps are control minus shock-condition means over $T21, T24, T27, T30$. Contrasts are computed before rounding.

	Shock			Recovery	
Setting	Δ_A	Δ_B	Δ_{amp}	A1–A2	B1–B2
DeepSeek / FS	1.75	8.77	+7.02	0.00	3.95
DeepSeek / PHYSICS	1.10	3.75	+2.65	0.08	4.15
GPT-5.4 / FS	1.75	3.51	+1.75	-2.63	5.26
GPT-5.2 / PHYSICS	2.14	6.70	+4.56	5.96	1.53

Moving beyond passive resource limits, we implement an active investment mechanism where agents can autonomously choose to purchase additional tokens prior to taking action, requiring them to recognize the strategic value of compute and make investments that shape future market outcomes. The *Token Flywheel* describes this potential investment feedback.

6.1 Purchasing and Investment

Derivers start with 2,048-token allowances and can spend wealth before acting to upgrade to 4,096, 6,144, or 8,192 tokens. Upgrades persist within an individual's lifetime but are not inherited. Team and individual rewards receive weights of 60% and 40%, respectively, while bidding, reproduction, and elimination remain active. Prices and configurations appear in Appendix G.

Across the four model–dataset settings, early token investment can give lineages a resource advantage for subsequent expansion. After founders exit, descendants continue purchasing additional capacity from the base allowance, sustaining investment across generations (Figure 5). Across all four settings, the descendants still alive at $T30$ in the final dominant lineage have positive aggregate net returns. Figure 7 traces lifetime returns for this fixed group alongside the founder's individual return. Full trajectories appear in Appendix G.2.

6.2 Purchase Restrictions and Lineage Outcomes

We restrict purchases by selected members of lineages that dominate the original runs. Paired conditions match the population and economic state before intervention. Only the target individual

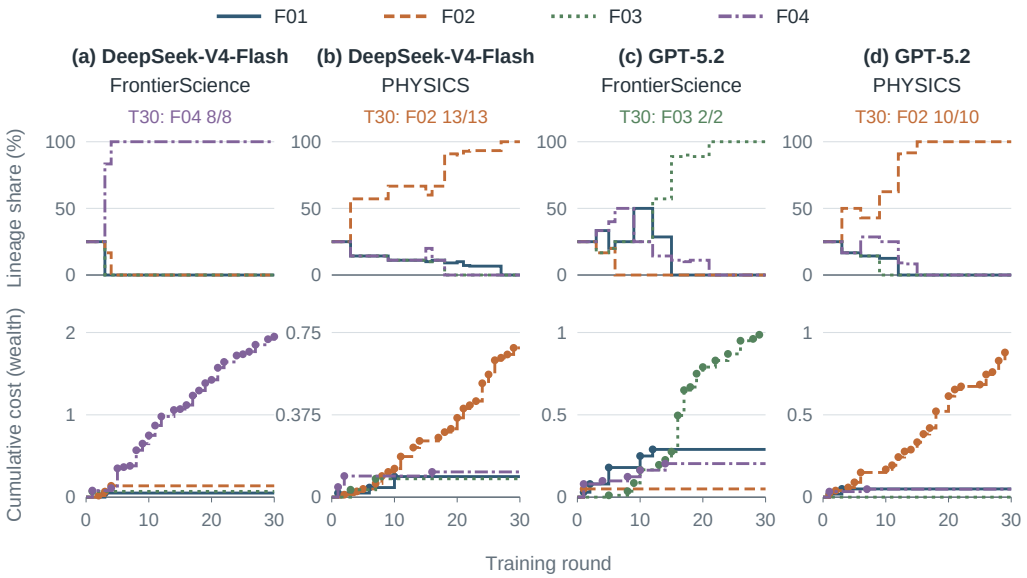

Figure 5: Purchasing and lineage dynamics across four configurations. Columns show configurations; rows show founder-lineage shares among surviving Derivers and cumulative purchase costs in wealth units. F01–F04 are local to each run. Dots mark charged purchases, including non-refunded charges; cost scales differ across columns. Shares are recorded after each task and its market updates.

Table 5: Individual-level purchase restrictions. Counts are target-lineage members/all surviving Derivers at $T30$. Lineage identifiers are local to each run.

Backbone	Lineage	Target	Window	Original	Restricted
FrontierScience					
DeepSeek-V4-Flash	F04	Founder	$T1$–$T3$	8/8	0/2
GPT-5.2	F03	D26	$T10$–$T12$	2/2	1/2
PHYSICS					
DeepSeek-V4-Flash	F02	Founder	$T2$–$T3$	13/13	0/11
GPT-5.2	F02	Founder	$T1$–$T3$	10/10	1/14

is restricted to 2,048 tokens; descendants remain free to purchase. Afterward, purchasing reopens without automatically granting a higher allowance.

The target lineage's final share decreases in all four comparisons (Table 5). On DeepSeek–FrontierScience, founder D7 retains its first action but cannot upgrade during $T1$–$T3$. F04 disappears by $T4$, whereas it occupies all eight final Deriver positions in the original run. On DeepSeek–PHYSICS, the restricted lineage disappears by $T18$, despite purchasing reopening at $T4$.

High-quality contributions can earn rewards that accumulate as wealth and fund larger token allowances. These resources can support further contributions and investment, forming a flywheel that promotes lineage expansion and dominance. However, greater lineage dominance does not necessarily improve team performance. For GPT-5.2, restricting purchases reduces the target lineage's final share in both settings, while training-task successes decrease from 20/30 to 18/30 on PHYSICS but increase from 13/30 to 16/30 on FrontierScience. Thus, a more dominant lineage can coexist with lower collective task performance. Detailed intervention trajectories appear in Appendix G.4.

6.3 Price Sensitivity

We vary purchase prices by factors of 0, 4, and 16; free upgrades still require an agent's choice. Mean post-decision quota decreases and inability to upgrade rises with price in all four configurations. Conditional no-upgrade rates also increase with price, indicating that agents more frequently forgo upgrades even when affordable options are available.

Lower prices tend to support larger Deriver populations and higher wealth shares, although this pattern is not uniform across configurations. Detailed resource, purchasing, and population measures appear in Appendix G.5. Together, these experiments connect wealth-funded access to lineage outcomes without implying that cheaper tokens or stronger lineage concentration always benefit the team.

7 Implications for Agent Economy

What should economic selection reward? An agent economy that produces more agents does not necessarily build a better problem-solving team. Our experiments show that while limiting token purchases successfully controls a lineage's growth, team performance varies across different task settings. This happens because of a core disconnect: individual agents compete and reproduce for private rewards, but solve problems as a team. Multi-agent system design must therefore align individual incentives with team success, rather than assuming that population growth equals better capability. This alignment does not mean banning market dominance or always favoring stronger models; weaker models with extra tokens are often more cost-effective. Instead, we should evaluate the final agent population by how well they perform tasks together using independent testing standards.

Who should receive the next tokens? The highest-scoring agent is not always the one that benefits most from extra compute. Similarly, accumulated wealth reflects past economic success, and Token Productivity (TP) measures recent efficiency; neither directly predicts how much the team will benefit from the next token allocation. This points to a better strategy: allocating resources based on an agent's marginal contribution to the final answer.

Crucially, this allocation must account for teamwork. For example, giving a Deriver more reasoning tokens only helps if downstream agents can actually use that extra reasoning to produce a better output. Future experiments should test this marginal-value approach under a fixed total compute budget, comparing it against allocation rules based on wealth, quality scores, or TP.

Can the economy revise its choices? Restoring a budget does not undo the population changes that occurred while it was restricted. Our recovery results raise a broader question: can an economy reconsider agents it previously eliminated? This matters both when an earlier resource disadvantage ends and when task requirements change. An agent that contributes little today may be useful on a different task tomorrow, but keeping it active also has a cost. Periodic reevaluation or limited re-entry could offer alternatives to permanent retention. Their value should be tested through performance after task shifts and the computation required to recover, rather than population diversity alone. The goal is not to preserve every lineage, but to avoid making useful capabilities costly or impossible to recover.

8 Conclusion

We studied how token resources shape competition in evolving agent economies. Across the reported experiments, advantages under matched token budgets do not reliably predict which model groups survive under unequal resources. Temporary budget reductions produce larger performance losses under economic evolution, while restoring budgets does not necessarily recover the unshocked trajectory. Paired purchase restrictions further show that token investment can support subsequent lineage expansion, without uniformly improving team performance. Token-productivity-based parent selection also improves mean checkpoint performance in the tested settings. Together, these results distinguish competitive success from model capability and collective task quality. Token allocation should therefore be designed and evaluated alongside reward and reproduction rules, rather than treated only as an inference setting. Establishing how these patterns generalize requires comparisons across economic mechanisms.

AI Use Statement

Generative AI tools were used solely to polish the language of the manuscript and improve readability. They were not used for any research tasks requiring disclosure under the conference's AI policy. All AI-assisted edits were reviewed by the authors to ensure that they preserved the intended meaning and technical accuracy. The authors take full responsibility for the final content of this work.

Ethics statement

This work studies how unequal computational resources can influence outcomes in multi-agent systems. We will report model access, costs, failure modes, and any potential implications for

manipulation or unfair allocation. The experiments do not involve human subjects or private personal data.

Reproducibility statement

The final version will document prompts, model versions, sampling parameters, token accounting, stopping rules, evaluation protocols, and statistical tests. Code and non-restricted experimental artifacts will be released in an anonymous supplement during review.

References

Justin Chih-Yao Chen, Swarnadeep Saha, and Mohit Bansal. Reconcile: Round-table conference improves reasoning via consensus among diverse llms, 2024. URL https://arxiv.org/abs/2309.13007.

Lingjiao Chen, Matei Zaharia, and James Zou. Frugalgpt: How to use large language models while reducing cost and improving performance, 2023a. URL https://arxiv.org/abs/2305.05176.

Weize Chen, Yusheng Su, Jingwei Zuo, Cheng Yang, Chenfei Yuan, Chi-Min Chan, Heyang Yu, Yaxi Lu, Yi-Hsin Hung, Chen Qian, Yujia Qin, Xin Cong, Ruobing Xie, Zhiyuan Liu, Maosong Sun, and Jie Zhou. Agentverse: Facilitating multi-agent collaboration and exploring emergent behaviors, 2023b. URL https://arxiv.org/abs/2308.10848.

Xingyu Chen, Jiahao Xu, Tian Liang, Zhiwei He, Jianhui Pang, Dian Yu, Linfeng Song, Qiuzhi Liu, Mengfei Zhou, Zhuosheng Zhang, Rui Wang, Zhaopeng Tu, Haitao Mi, and Dong Yu. Do not think that much for 2+3=? on the overthinking of o1-like llms, 2025. URL https://arxiv.org/abs/2412.21187.

Yufan Dang, Chen Qian, Xueheng Luo, Jingru Fan, Zihao Xie, Ruijie Shi, Weize Chen, Cheng Yang, Xiaoyin Che, Ye Tian, Xuantang Xiong, Lei Han, Zhiyuan Liu, and Maosong Sun. Multi-agent collaboration via evolving orchestration, 2025. URL https://arxiv.org/abs/2505.19591.

DeepSeek-AI, Anyi Xu, Bangcai Lin, Bing Xue, and etc. Deepseek-v4: Towards highly efficient million-token context intelligence, 2026. URL https://arxiv.org/abs/2606.19348.

Yilun Du, Shuang Li, Antonio Torralba, Joshua B. Tenenbaum, and Igor Mordatch. Improving factuality and reasoning in language models through multiagent debate, 2023. URL https://arxiv.org/abs/2305.14325.

Kaiyue Feng, Yilun Zhao, Yixin Liu, Tianyu Yang, Chen Zhao, John Sous, and Arman Cohan. Physics: Benchmarking foundation models on university-level physics problem solving, 2025. URL https://arxiv.org/abs/2503.21821.

Mor Geva, Daniel Khashabi, Elad Segal, Tushar Khot, Dan Roth, and Jonathan Berant. Did aristotle use a laptop? a question answering benchmark with implicit reasoning strategies, 2021. URL https://arxiv.org/abs/2101.02235.

Tingxu Han, Zhenting Wang, Chunrong Fang, Shiyu Zhao, Shiqing Ma, and Zhenyu Chen. Token-budget-aware llm reasoning, 2025. URL https://arxiv.org/abs/2412.18547.

Sirui Hong, Mingchen Zhuge, Jiaqi Chen, Xiawu Zheng, Yuheng Cheng, Ceyao Zhang, Jinlin Wang, Zili Wang, Steven Ka Shing Yau, Zijuan Lin, Liyang Zhou, Chenyu Ran, Lingfeng Xiao, Chenglin Wu, and Jürgen Schmidhuber. Metagpt: Meta programming for a multi-agent collaborative framework, 2024. URL https://arxiv.org/abs/2308.00352.

Qitian Jason Hu, Jacob Bieker, Xiuyu Li, Nan Jiang, Benjamin Keigwin, Gaurav Ranganath, Kurt Keutzer, and Shriyash Kaustubh Upadhyay. Routerbench: A benchmark for multi-llm routing system, 2024. URL https://arxiv.org/abs/2403.12031.

Shengran Hu, Cong Lu, and Jeff Clune. Automated design of agentic systems, 2025. URL https://arxiv.org/abs/2408.08435.

Sayash Kapoor, Benedikt Stroebl, Zachary S. Siegel, Nitya Nadgir, and Arvind Narayanan. Ai agents that matter, 2024. URL https://arxiv.org/abs/2407.01502.

Yubin Kim, Ken Gu, Chanwoo Park, Chunjong Park, Samuel Schmidgall, A. Ali Heydari, Yao Yan, Zhihan Zhang, Yuchen Zhuang, Yun Liu, Mark Malhotra, Paul Pu Liang, Hae Won Park, Yuzhe Yang, Xuhai Xu, Yilun Du, Shwetak Patel, Tim Althoff, Daniel McDuff, and Xin Liu. Towards a science of scaling agent systems, 2026. URL https://arxiv.org/abs/2512.08296.

Guohao Li, Hasan Abed Al Kader Hammoud, Hani Itani, Dmitrii Khizbullin, and Bernard Ghanem. Camel: Communicative agents for "mind" exploration of large language model society, 2023. URL https://arxiv.org/abs/2303.17760.

Junyou Li, Qin Zhang, Yangbin Yu, Qiang Fu, and Deheng Ye. More agents is all you need, 2024a. URL https://arxiv.org/abs/2402.05120.

Nian Li, Chen Gao, Mingyu Li, Yong Li, and Qingmin Liao. Econagent: Large language model-empowered agents for simulating macroeconomic activities, 2024b. URL https://arxiv.org/abs/2310.10436.

Yunxuan Li, Yibing Du, Jiageng Zhang, Le Hou, Peter Grabowski, Yeqing Li, and Eugene Ie. Improving multi-agent debate with sparse communication topology, 2024c. URL https://arxiv.org/abs/2406.11776.

Zijun Liu, Yanzhe Zhang, Peng Li, Yang Liu, and Diyi Yang. A dynamic llm-powered agent network for task-oriented agent collaboration, 2024. URL https://arxiv.org/abs/2310.02170.

Haotian Luo, Li Shen, Haiying He, Yibo Wang, Shiwei Liu, Wei Li, Naiqiang Tan, Xiaochun Cao, and Dacheng Tao. O1-pruner: Length-harmonizing fine-tuning for o1-like reasoning pruning, 2025. URL https://arxiv.org/abs/2501.12570.

Niklas Muennighoff, Zitong Yang, Weijia Shi, Xiang Lisa Li, Li Fei-Fei, Hannaneh Hajishirzi, Luke Zettlemoyer, Percy Liang, Emmanuel Candès, and Tatsunori Hashimoto. s1: Simple test-time scaling, 2025. URL https://arxiv.org/abs/2501.19393.

Isaac Ong, Amjad Almahairi, Vincent Wu, Wei-Lin Chiang, Tianhao Wu, Joseph E. Gonzalez, M Waleed Kadous, and Ion Stoica. Routellm: Learning to route llms with preference data, 2025. URL https://arxiv.org/abs/2406.18665.

Joon Sung Park, Joseph O'Brien, Carrie Jun Cai, Meredith Ringel Morris, Percy Liang, and Michael S Bernstein. Generative agents: Interactive simulacra of human behavior. In *Proceedings of the 36th annual acm symposium on user interface software and technology*, pp. 1–22, 2023.

Zhenting Qi, Huangyuan Su, Ao Qu, Chenyu Wang, Yu Yao, Han Zheng, Kushal Chattopadhyay, Guowei Xu, Zihan Wang, Weirui Ye, Vijay Janapa Reddi, Ju Li, Paul Pu Liang, Himabindu Lakkaraju, Sham Kakade, and Yilun Du. Economy of minds: Emerging multi-agent intelligence with economic interactions, 2026. URL https://arxiv.org/abs/2606.02859.

Chen Qian, Wei Liu, Hongzhang Liu, Nuo Chen, Yufan Dang, Jiahao Li, Cheng Yang, Weize Chen, Yusheng Su, Xin Cong, Juyuan Xu, Dahai Li, Zhiyuan Liu, and Maosong Sun. Chatdev: Communicative agents for software development, 2024. URL https://arxiv.org/abs/2307.07924.

Charlie Snell, Jaehoon Lee, Kelvin Xu, and Aviral Kumar. Scaling llm test-time compute optimally can be more effective than scaling model parameters, 2024. URL https://arxiv.org/abs/2408.03314.

Junlin Wang, Jue Wang, Ben Athiwaratkun, Ce Zhang, and James Zou. Mixture-of-agents enhances large language model capabilities, 2024. URL https://arxiv.org/abs/2406.04692.

Miles Wang et al. Frontierscience: Evaluating ai's ability to perform expert-level scientific tasks, 2026. URL https://arxiv.org/abs/2601.21165.

Xuezhi Wang, Jason Wei, Dale Schuurmans, Quoc Le, Ed Chi, Sharan Narang, Aakanksha Chowdhery, and Denny Zhou. Self-consistency improves chain of thought reasoning in language models, 2023. URL https://arxiv.org/abs/2203.11171.

Zhaotian Weng, Antonis Antoniades, Deepak Nathani, Zhen Zhang, Xiao Pu, and Xin Eric Wang. Group-evolving agents: Open-ended self-improvement via experience sharing, 2026. URL https://arxiv.org/abs/2602.04837.

Qingyun Wu, Gagan Bansal, Jieyu Zhang, Yiran Wu, Beibin Li, Erkang Zhu, Li Jiang, Xiaoyun Zhang, Shaokun Zhang, Jiale Liu, Ahmed Hassan Awadallah, Ryen W White, Doug Burger, and Chi Wang. Autogen: Enabling next-gen llm applications via multi-agent conversation, 2023. URL https://arxiv.org/abs/2308.08155.

Yingxuan Yang, Ying Wen, Jun Wang, and Weinan Zhang. Agent exchange: Shaping the future of ai agent economics, 2025. URL https://arxiv.org/abs/2507.03904.

Shunyu Yao, Dian Yu, Jeffrey Zhao, Izhak Shafran, Thomas L. Griffiths, Yuan Cao, and Karthik Narasimhan. Tree of thoughts: Deliberate problem solving with large language models, 2023. URL https://arxiv.org/abs/2305.10601.

Gokul Yenduri, Ramalingam M, Chemmalar Selvi G, Supriya Y, Gautam Srivastava, Praveen Kumar Reddy Maddikunta, Deepti Raj G, Rutvij H Jhaveri, Prabadevi B, Weizheng Wang, Athanasios V. Vasilakos, and Thippa Reddy Gadekallu. Generative pre-trained transformer: A comprehensive review on enabling technologies, potential applications, emerging challenges, and future directions, 2023. URL https://arxiv.org/abs/2305.10435.

Siyu Yuan, Kaitao Song, Jiangjie Chen, Xu Tan, Dongsheng Li, and Deqing Yang. Evoagent: Towards automatic multi-agent generation via evolutionary algorithms, 2025. URL https://arxiv.org/abs/2406.14228.

Guibin Zhang, Yanwei Yue, Zhixun Li, Sukwon Yun, Guancheng Wan, Kun Wang, Dawei Cheng, Jeffrey Xu Yu, and Tianlong Chen. Cut the crap: An economical communication pipeline for llm-based multi-agent systems, 2024. URL https://arxiv.org/abs/2410.02506.

Jenny Zhang, Shengran Hu, Cong Lu, Robert Lange, and Jeff Clune. Darwin godel machine: Open-ended evolution of self-improving agents, 2026. URL https://arxiv.org/abs/2505.22954.

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, 2025. URL https://arxiv.org/abs/2410.10762.

Qinlin Zhao, Jindong Wang, Yixuan Zhang, Yiqiao Jin, Kaijie Zhu, Hao Chen, and Xing Xie. Competeai: Understanding the competition dynamics in large language model-based agents, 2024. URL https://arxiv.org/abs/2310.17512.

Stephan Zheng, Alexander Trott, Sunil Srinivasa, David C. Parkes, and Richard Socher. The ai economist: Optimal economic policy design via two-level deep reinforcement learning, 2021. URL https://arxiv.org/abs/2108.02755.

Mingchen Zhuge, Wenyi Wang, Louis Kirsch, Francesco Faccio, Dmitrii Khizbullin, and Jürgen Schmidhuber. Language agents as optimizable graphs, 2024. URL https://arxiv.org/abs/2402.16823.

A Datasets and Filtering

FrontierScience-Research. We use the FrontierScience-Research training and evaluation splits (Wang et al., 2026) provided by EOM (Qi et al., 2026). The training file contains 40 tasks; our economic experiments use its first 30 tasks for evolution and all 19 tasks in the separate evaluation file for checkpoint assessment.

PHYSICS. We select text-only problems from PHYSICS (Feng et al., 2025) with multi-step reference derivations and one or two final answers. The fixed evaluation set contains 20 problems from the official test split with 11–15 reference-solution display-math blocks. The 30 evolution tasks come from the official `eval` split: 29 eligible problems with 9–15 such blocks and the shortest eligible eight-block problem. The evolution and evaluation sets do not overlap.

B Related Work

Multi-Agent Collaboration and Debate LLM-based multi-agent systems use communication and role specialization to improve task performance. Debate and consensus methods combine reasoning from multiple agents (Du et al., 2023; Chen et al., 2024), while Mixture-of-Agents integrates model outputs through a layered architecture (Wang et al., 2024). CAMEL and AutoGen support role-based and configurable conversations (Li et al., 2023; Wu et al., 2023); MetaGPT and ChatDev organize specialized agents into structured workflows (Hong et al., 2024; Qian et al., 2024). Beyond predefined teams, AgentVerse and DyLAN adapt agent composition and interactions (Chen et al., 2023b; Liu et al., 2024). GPTSwarm optimizes prompts and communication graphs, while AFlow and Evolving Orchestration automate workflow design and agent coordination (Zhuge et al., 2024; Zhang et al., 2025; Dang et al., 2025). Studies of agent count, sparse communication, and coordination overhead further examine when additional agents improve performance and at what cost (Li et al., 2024a;c; Kim et al., 2026). We study a different consequence of resource allocation: how token differences affect agents' continued participation and lineage survival across tasks.

Agent Economies and Population Dynamics Agent-based simulations connect individual decisions to collective outcomes. Generative Agents studies emergent social behavior (Park et al., 2023), while EconAgent models macroeconomic activity through LLM agents' work and consumption decisions (Li et al., 2024b). Using reinforcement learning rather than LLMs, the AI Economist studies how economic agents and a tax-setting planner adapt to one another (Zheng et al., 2021). CompeteAI examines competition in simulated markets (Zhao et al., 2024), and Agent Exchange proposes auction-based allocation of tasks to agent services (Yang et al., 2025). Economy of Minds (EOM) connects task contributions to wealth, survival, and reproduction (Qi et al., 2026). We use this setting to study how token access shapes population dynamics.

A related but distinct direction uses search and evolution to improve agents. ADAS searches over code-defined agent designs (Hu et al., 2025), and EvoAgent applies evolutionary operators to generate agent configurations (Yuan et al., 2025). The Darwin Gödel Machine explores self-modifying coding agents (Zhang et al., 2026), while Group-Evolving Agents supports self-improvement through shared experience (Weng et al., 2026). These methods seek better agent designs. Our question concerns selection within an economy: how persistent resource differences, temporary shocks, and token investment affect which lineages survive.

Compute Allocation and Resource-Constrained Agents Additional inference computation can improve reasoning, but its value depends on how it is used. Self-consistency aggregates sampled reasoning paths (Wang et al., 2023), and Tree of Thoughts searches over intermediate reasoning steps (Yao et al., 2023). Other work studies problem-dependent compute allocation (Snell et al., 2024) and direct control of reasoning budgets (Muennighoff et al., 2025). Conversely, research on overthinking identifies redundant computation (Chen et al., 2025). Token-Budget-Aware Reasoning and O1-Pruner seek to reduce reasoning length while preserving accuracy (Han et al., 2025; Luo et al., 2025).

At the system level, FrugalGPT and RouteLLM balance performance and cost through model cascades and routing (Chen et al., 2023a; Ong et al., 2025), with RouterBench providing a framework for evaluating routing systems (Hu et al., 2024). AgentPrune reduces redundant multi-agent communication (Zhang et al., 2024), and cost-aware agent evaluation emphasizes reporting expense alongside accuracy (Kapoor et al., 2024). These studies address the effective use and evaluation of computation. We examine how access to computation changes over time: task rewards affect wealth and reproduction, and wealth-funded purchases can expand future token budgets. This feedback makes resource allocation relevant not only to current task performance but also to future population composition.

Table 6: DeepSeek FrontierScience: restricting founder D7's purchases in $T1$–$T3$ while preserving its initial action opportunity. Net contribution return is committed individual and team rewards minus token-purchase charges, excluding rent and bidding transfers. The test count pools the same 19 fixed questions across ten checkpoints.

Measure	Original	Restriction
D7 quota at $T1$	$2048 \to 6144 \to 8192$	2048
D7 net contribution return	$+0.448$	-0.436
D7 exit	$T18$	$T3$
Observed descendants	40	1
Final F04 Derivers	$8/8$	$0/2$
Final winning lineage	F04	F03 $(2/2)$
Pooled checkpoint passes (≥ 0.7)	$34/190$ (17.9%)	$32/190$ (16.8%)

Table 7: DeepSeek PHYSICS: F02 lineage after restricting founder D5's purchases in $T2$–$T3$. Wealth shares are within the surviving Deriver role. Purchasing becomes available again at $T4$.

	Original		Restriction	
Round	F02/Derivers	Wealth (%)	F02/Derivers	Wealth (%)
$T3$	$4/7$	57.8	$1/6$	20.0
$T9$	$6/9$	76.5	$1/13$	4.5
$T18$	$10/11$	87.9	$0/11$	0.0
$T30$	$13/13$	100.0	$0/11$	0.0

C Additional Token-Purchasing Results

C.1 Early-Purchase Counterfactuals

The restrictions apply to the listed individual, not to its descendants. Table 6 retains the FrontierScience event sequence, and Table 7 gives the PHYSICS lineage and wealth snapshots.

D Additional Results for Resource-Dependent Competition

D.1 Population, Wealth, and Actions

Table 8 supplements the population results in Section 4 with wealth and public action shares for the 8,192/2,048-token conditions. Each model group initially contains two founder lineages. The reported shares aggregate these lineages within the Deriver role and do not describe the entire market.

DeepSeek–FrontierScience. Flash produces seven descendants and Pro produces two. Their cumulative public action counts are similar (39 versus 37), but Flash finishes with two Derivers and 1.5723 wealth, while Pro has no surviving agents. Pro temporarily holds 75% of the two groups' population and 79.9% of their wealth at $T12$; its last descendant is eliminated at $T27$. These events illustrate a change in competitive position over time.

DeepSeek–PHYSICS. Both groups produce three descendants. Flash reaches four Derivers at $T6$, while Pro returns to two at $T15$ before disappearing at $T23$. Flash finishes with one Deriver and 0.3803 wealth. Cumulative public actions are 35 for Flash and 17 for Pro.

GPT–FrontierScience. GPT-5.2 reaches three Derivers at $T3$, falls to one with 0.1 wealth at $T12$, and disappears at $T15$. GPT-5.1 finishes with one Deriver and 0.6060 wealth. Although this agent holds all remaining Deriver wealth, it accounts for only 10.3% of total market wealth and 5.0% of the market population. Dominance within a role therefore need not imply dominance over the entire economy.

Table 8: Final outcomes with 8,192 tokens for weaker-model Derivers and 2,048 for stronger-model Derivers. Weak/strong denotes Flash/Pro or GPT-5.1/GPT-5.2. Wealth shares are measured at $T30$; action shares use cumulative public actions. Share denominators include the two competing Deriver model groups.

Family	Dataset	Final agents weak/strong	Weak wealth share (%)	Weak action share (%)
DeepSeek	FrontierScience	2/0	100.0	51.3
	PHYSICS	1/0	100.0	67.3
GPT	FrontierScience	1/0	100.0	78.0
	PHYSICS	10/1	94.4	80.4

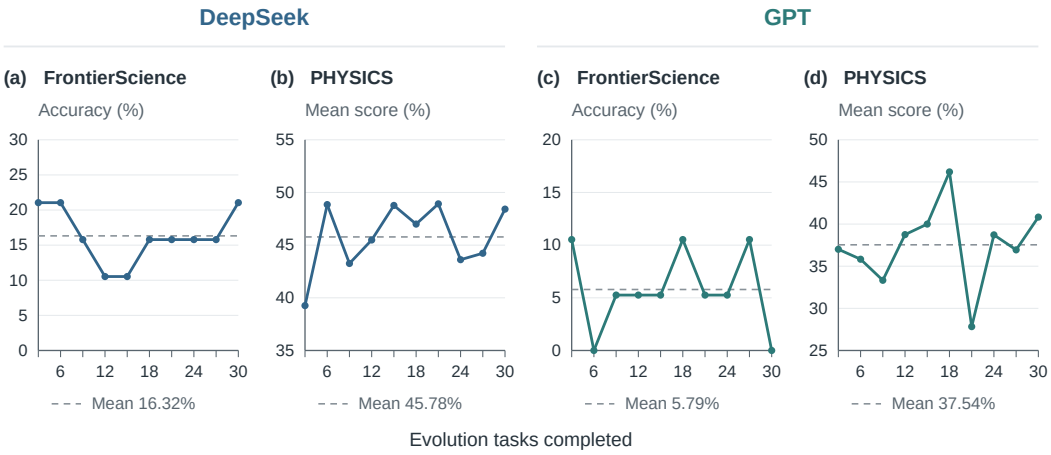

Figure 6: Held-out performance under wealth-based parent selection and unequal token budgets: (a,b) DeepSeek and (c,d) GPT. FrontierScience reports accuracy (a,c); PHYSICS reports mean score (b,d). Dashed lines indicate ten-checkpoint means; panels use different vertical scales.

GPT–PHYSICS. The GPT-5.1/GPT-5.2 populations move from 3/2 at $T3$ to 3/3 at $T6$ and 9/1 at $T18$. Both model groups remain represented at $T30$, with populations of 10/1, wealth of 3.3667/0.2000, and cumulative public actions of 41/10. This setting exhibits a large population imbalance, rather than complete disappearance of the stronger model group.

D.2 Held-Out Performance Trajectories

Figure 6 reports frozen held-out evaluations for these runs. FrontierScience uses thresholded accuracy, whereas PHYSICS uses mean per-question score. Dashed lines indicate averages over ten checkpoints.

On GPT–PHYSICS, the score changes from 46.19% at $T18$ to 27.83% at $T21$ and 40.83% at $T30$, despite the weaker-model group's large population share.

D.3 Detailed Outcomes across Token Budgets

Table 9 provides wealth and action counts for the budget comparisons in Section 4.2. The stronger model receives 2,048 tokens throughout. Counts and wealth are final values; public actions are cumulative over the run.

Table 9: Detailed budget-comparison outcomes. Entries give weaker/stronger values: Flash/Pro for DeepSeek and GPT-5.1/GPT-5.2 for GPT. Wealth is measured in internal market units.

Family	Dataset	Weak-model budget	Final agents	Final wealth	Actions
DeepSeek	FrontierScience	2,048	0/1	0/0.2335	21/59
		6,144	4/0	1.3133/0	72/11
		8,192	2/0	1.5723/0	39/37
DeepSeek	PHYSICS	2,048	0/13	0/4.4127	13/47
		6,144	0/8	0/2.3000	11/44
		8,192	1/0	0.3803/0	35/17
GPT	FrontierScience	2,048	0/6	0/2.6940	16/50
		6,144	2/0	0.7667/0	62/14
		8,192	1/0	0.6060/0	46/13
GPT	PHYSICS	2,048	0/3	0/1.8889	8/46
		6,144	1/1	1.1167/0.3000	26/22
		8,192	10/1	3.3667/0.2000	41/10

E Token Productivity

E.1 Quality and Token-Usage Estimates

TP uses at most the three most recent contributions available for agent i at selection time t. Let $n_i(t) \leq 3$ be the number of available contributions, q_{ik} the quality score assessed by GPT-5.2, and u_{ik} the actual generation-token consumption for contribution k. Consumption includes both private reasoning and public output, rather than the assigned token allowance.

For $n_i(t) > 0$, we compute

$$\widetilde{q}_i(t) = \frac{0.5 + \sum_{k=1}^{n_i(t)} q_{ik}}{1 + n_i(t)}, \tag{4}$$

$$\widetilde{u}_i(t) = \frac{1}{n_i(t)} \sum_{k=1}^{n_i(t)} u_{ik}. \tag{5}$$

These estimates enter Equation 2. Agents without a contribution history receive zero TP.

The quality estimate adds one prior observation with score 0.5. Writing the unadjusted mean as $\overline{q}_i(t)$ gives

$$\widetilde{q}_i(t) - 0.5 = \frac{n_i(t)}{n_i(t) + 1}\big(\overline{q}_i(t) - 0.5\big). \tag{6}$$

Thus, the excess-quality estimate is shrunk toward zero, with stronger shrinkage for shorter histories. The 0.5 baseline is a contribution-level reference and is distinct from the 0.7 threshold used to classify FrontierScience final answers.

E.2 Resource Normalization

TP divides positive excess quality by $\max(1, \widetilde{u}_i(t)/2048)$. For average consumption at or below 2,048 tokens, the denominator is one: reducing consumption further does not increase TP at fixed quality. Above 2,048 tokens, higher consumption lowers TP unless accompanied by sufficient improvement in quality.

TP is therefore a baseline-adjusted selection score, not a direct estimate of intrinsic model capability or an unrestricted quality-per-token ratio. Its denominator uses actual generation consumption; it does not account for monetary prices or other computational costs.

Table 10: Mean performance (%) over ten checkpoints, corresponding to Figure 4. FS denotes FrontierScience accuracy; PHYSICS reports mean score. Δ is TP minus Wealth in percentage points, computed before rounding.

Family	Dataset	Wealth	TP	Δ
DeepSeek	FS	16.32	21.05	+4.7
	PHYSICS	45.78	51.38	+5.6
GPT	FS	5.79	7.89	+2.1
	PHYSICS	37.54	39.13	+1.6

Table 11: Held-out performance (%) across intervention phases. FrontierScience reports thresholded accuracy over 19 tasks; PHYSICS reports mean score over 20 tasks. Window means and $T30$ results are distinct summaries.

Condition	Pre-shock	Shock	Recovery	$T30$
DeepSeek-V4-Flash / FrontierScience				
A1	15.79	17.54	14.47	15.79
A2	15.79	15.79	14.47	15.79
B1	15.79	19.30	19.74	26.32
B2	15.79	10.53	15.79	21.05
DeepSeek-V4-Flash / PHYSICS				
A1	45.62	45.50	44.24	36.73
A2	45.62	44.40	44.16	44.67
B1	45.62	45.64	44.49	41.01
B2	45.62	41.89	40.34	47.49
GPT-5.4 / FrontierScience				
A1	21.05	14.04	13.16	15.79
A2	21.05	12.28	15.79	21.05
B1	21.05	17.54	15.79	15.79
B2	21.05	14.04	10.53	10.53
GPT-5.2 / PHYSICS				
A1	41.82	42.20	41.13	36.42
A2	41.82	40.06	35.17	30.60
B1	41.82	46.00	43.04	47.92
B2	41.82	39.30	41.50	42.11

E.3 Selection and Evaluation

The TP condition uses this score instead of accumulated wealth as the criterion for selecting Deriver parents. Table 10 reports mean held-out performance across ten checkpoints for each condition. These comparisons evaluate the complete selection variant, including quality scoring, the recent-history window, smoothing, and resource normalization.

F Additional Results for Temporary Resource Shocks

F.1 Evaluation Windows and Performance

All four conditions share their history through $T10$. Pre-shock means use checkpoints $T3, T6, T9$; shock means use $T12, T15, T18$; and recovery means use $T21, T24, T27, T30$. Evaluations repeatedly use the fixed held-out questions. Table 11 reports the window means and final-checkpoint values underlying Table 3.

The recovery comparison uses the contemporaneous unchanged-budget control rather than only the pre-shock baseline. This distinction matters because the control market continues to evolve. Window averages also differ from endpoint comparisons: on DeepSeek–PHYSICS, B2 has lower mean recovery performance but higher performance at $T30$.

Table 12: Allowance tiers and cumulative base prices in internal market wealth units. Upgrades pay the difference between tier prices.

Allowance b	2,048	4,096	6,144	8,192
Cumulative price $C(b)$	0	0.015	0.040	0.080

F.2 Population and Wealth Snapshots

Table 4 reports end-of-task states at $T10$, $T20$, and $T30$. Population and wealth are aggregated over the Deriver role, not individual founder lineages. The table includes market totals so that absolute outcomes can be distinguished from role shares.

In both FrontierScience settings and GPT–PHYSICS, B2 has more Derivers than B1 at $T20$ but fewer at $T30$; Deriver wealth follows the same ordering reversal.

DeepSeek–PHYSICS shows the opposite final ordering: B2 ends with two Derivers and 0.66 wealth, compared with one and 0.33 in B1. This ecological advantage coexists with lower average recovery performance. On GPT–PHYSICS, B2's lower recovery-window mean (41.50% versus 43.04%) accompanies its lower final Deriver population and wealth.

G Token Purchasing: Settings and Additional Results

G.1 Purchasing Mechanism

Purchasing experiments use DeepSeek-V4-Flash and GPT-5.2 on FrontierScience and PHYSICS. Each run starts with four Derivers and one agent in each of the Literature, Planner, Verifier, and Answer roles. All interacting agents within a run share a backbone; GPT-5.2 provides LLM-based judgments. Runs follow the 30-task evolution horizon in Section 3.

A purchase raises the allowance available for subsequent contributions, rather than adding tokens to a consumable lifetime balance. Given cumulative base prices $C(b)$ (Table 12), upgrading from b to b' costs

$$\mathrm{Cost}_{i,t}(b \to b') = \lambda_{i,t}[C(b') - C(b)], \qquad \lambda_{i,t} \in \{0.75, 1, 1.25\}. \tag{7}$$

The multiplier is sampled for each purchase quote. Costs are deducted from wealth. Descendants start at 2,048 tokens regardless of the parent's purchased allowance.

Team rewards based on final answers and individual rewards based on Deriver contributions are normalized, centered, and weighted by 60% and 40%, respectively. Both enter wealth, while purchases incur deductions. Bidding, reproduction, and elimination remain active.

F01–F04 identify founder lineages locally within each run. Individual labels such as D7 and D26 are agent identifiers, not generation numbers; parent links determine ancestry.

G.2 Investment and Lineage Trajectories

Table 13 summarizes four model–dataset configurations. Fixation refers to all surviving Derivers belonging to one founder lineage, not to control of the entire market. For example, GPT-5.2's F02 on PHYSICS ends with ten of ten Derivers but ten of twenty total agents.

On DeepSeek–FrontierScience, F04's founder upgrades to 8,192 tokens at $T1$. Its first three committed contributions score 0.62, 0.85, and 0.85, yielding $+0.656$ in individual rewards. F04 reaches five of six Derivers at $T3$ and seven of seven at $T4$. The founder exits at $T18$, but descendants continue purchasing.

The two PHYSICS examples reach fixation later. DeepSeek's F02 occupies all thirteen final Deriver positions, and GPT-5.2's F02 occupies all ten. A different DeepSeek–PHYSICS founder reaches the highest allowance by $T2$ but its lineage disappears around $T18$. Thus, early upgrading is not sufficient for persistence.

Figure 5 reports lineage shares and actual purchase expenditure in all four settings. For GPT-5.2 on FrontierScience, F03 reaches fixation by $T21$ and occupies both surviving Deriver positions at

Table 13: Purchasing and lineage outcomes across four model–dataset configurations. "Fixed by" denotes a recorded task showing fixation, not necessarily its exact onset. Final counts are lineage members/all Derivers at $T30$.

Dataset	Backbone	Lineage	Founder purchases	Fixed by	Final count
FrontierScience	DeepSeek-V4-Flash	F04	$T1$	$T4$	$8/8$
FrontierScience	GPT-5.2	F03	$T5$, $T8$, $T9$	$T21$	$2/2$
PHYSICS	DeepSeek-V4-Flash	F02	$T2$, $T9$, $T13$	$T27$	$13/13$
PHYSICS	GPT-5.2	F02	$T1$, $T2$, $T6$	$T15$	$10/10$

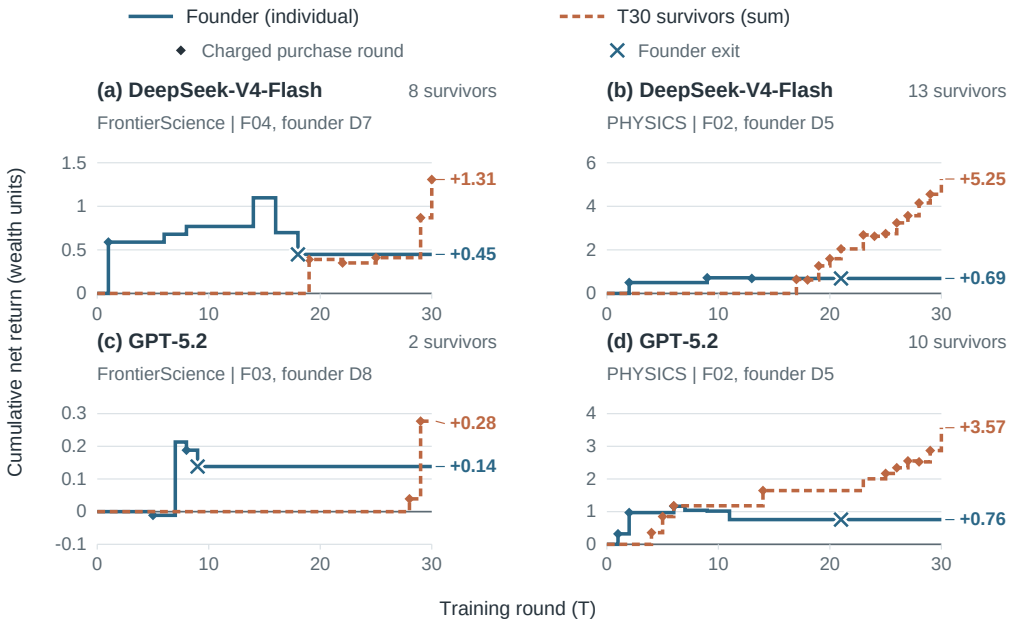

Figure 7: Founder and surviving-descendant returns across four settings. The dashed curve sums returns from birth for all descendants present at $T30$, with membership fixed across the plotted rounds. Net return accumulates personal and team rewards minus charged token-purchase costs, excluding rent and bidding transfers. F labels identify lineages; D labels identify individual Deriver founders. Diamonds mark charged purchase rounds and crosses founder exits; founder returns remain constant afterward. Endpoint labels give $T30$ returns. Vertical scales differ across panels.

$T30$. Figure 7 compares each founder with all descendants still alive at $T30$ in the lineage occupying all final Deriver positions in that run. These agents serve the Deriver role, producing reasoning contributions. The descendant curve sums lifetime returns for this fixed group of $T30$ survivors. In all four runs, these descendants continue purchasing after the founder exits and finish with positive aggregate net returns. Appendix G.3 reports the lineage accounts by final survival status. These trajectories document purchasing and expansion, but do not independently identify the effect of purchasing on fixation.

G.3 Descendant Returns by Final Survival Status

Table 14 accounts for all descendants of the target lineages in Figure 7. All members are Derivers; founders are accounted for separately. Descendants are grouped by survival at $T30$, with group membership held fixed when tracing their histories. All members of each group are retained regardless of return. Net return is committed individual rewards plus actual team payouts minus nonrefunded

Table 14: Complete descendant accounting at $T30$ for the four target lineages. Counts exclude the founder. Net returns are summed over members in wealth units and rounded to three decimals.

	$T30$ survivors		Exited		All descendants	
Configuration	n	Net return	n	Net return	n	Net return
DeepSeek-V4-Flash FrontierScience (F04)	8	$+1.308$	32	-7.448	40	-6.141
DeepSeek-V4-Flash PHYSICS (F02)	13	$+5.253$	9	$+2.090$	22	$+7.343$
GPT-5.2 FrontierScience (F03)	2	$+0.277$	13	-1.765	15	-1.488
GPT-5.2 PHYSICS (F02)	10	$+3.570$	17	$+1.037$	27	$+4.607$

token-purchase charges, excluding initial endowments, rent, bidding, and reproduction transfers. Exited members retain their historical returns in the totals.

G.4 Paired Purchase Restrictions

The population and economic state are matched before intervention. Only the listed individual is restricted to the base allowance during the window in Table 5. Other agents, including descendants, retain purchase access. After the window, the target may purchase again but does not automatically receive a higher allowance.

On DeepSeek–FrontierScience, D7 remains activated, wins the first auction, and contributes first. Restricting upgrades advances its exit from $T18$ to $T3$; F04 disappears by $T4$ rather than occupying all eight final Deriver positions. The initial action opportunity is therefore preserved, although subsequent trajectories diverge.

On DeepSeek–PHYSICS, F02 holds six of nine Deriver positions at $T9$ in the original run, compared with one of thirteen under restriction. It disappears by $T18$ despite purchasing reopening at $T4$.

In the GPT runs, target lineages remain represented but lose exclusivity. F03 on FrontierScience ends with one of two Derivers rather than two of two; F02 on PHYSICS ends with one of fourteen rather than ten of ten. Figure 8 shows the paired population trajectories.

G.5 Price Sensitivity and Purchasing Measures

Price experiments multiply Equation 7 by $\rho \in \{0, 4, 16\}$. Allowance tiers, quote multipliers, reward weights, and non-inheritance remain unchanged. Free upgrades still require an agent's choice. These runs use DeepSeek-V4-Flash and GPT-5.2 on FrontierScience and Yale (PHYSICS) for 30 tasks.

Mean post-decision quota is the mean private-plus-public capacity after each Deriver purchase decision during training, including decisions without an upgrade, with equal weight per decision. Repeated decisions by the same agent are counted separately; initial states and agents that have not yet acted do not contribute separate observations. Initial and newborn Derivers receive 2,048 tokens, but their first recorded post-decision quota can already be 8,192. Inability to upgrade uses all purchase decisions as its denominator; conditional no-upgrade uses decisions with at least one affordable upgrade. Final counts and wealth shares use all surviving agents and their total wealth after task 30 as denominators.

Mean post-decision quota falls and inability to upgrade rises with price in all four configurations. For DeepSeek–FrontierScience, conditional no-upgrade rates are similar under the two paid conditions (24.32% and 25.00%); for DeepSeek–PHYSICS, they rise from 5.00% to 64.86%. For GPT, conditional no-upgrade reaches 90.20% on FrontierScience and 90.91% on PHYSICS at $16\times$.

Final Deriver counts and wealth shares decline with price in both DeepSeek configurations. GPT–FrontierScience instead has its largest final Deriver count and wealth share at $4\times$, while GPT–

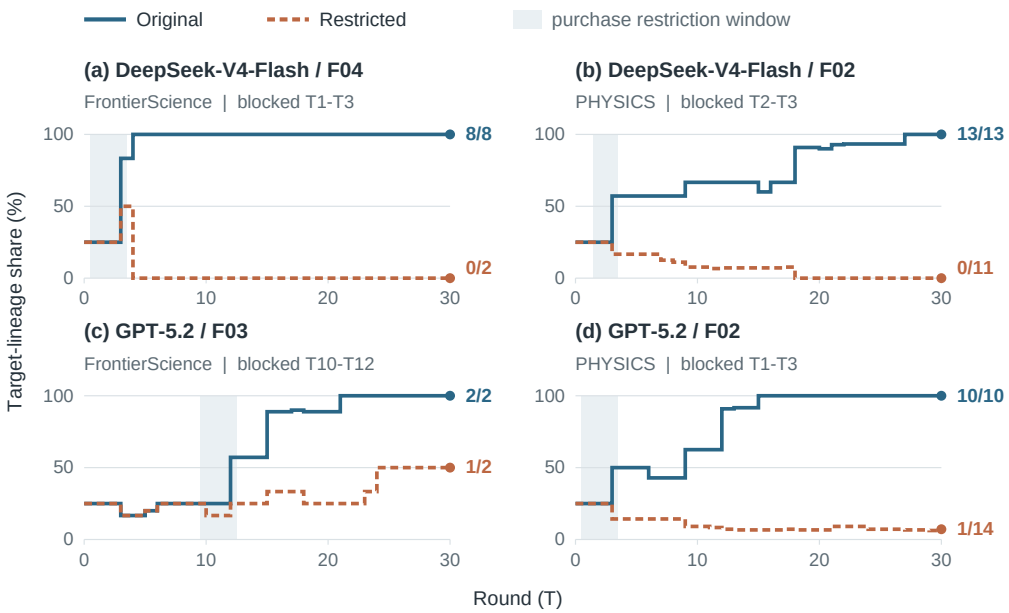

Figure 8: Target-lineage trajectories under early-purchase restrictions. Step lines show the target lineage's share of surviving Derivers after each task in the original and restricted conditions. Shading marks the individual-level restriction window; purchasing reopens afterward. Endpoint labels give target-lineage/total Deriver counts at $T30$.

PHYSICS retains 11 Derivers out of 20 agents at all three prices. These are role-level outcomes, unlike the founder-lineage comparisons in Section 6.2.

Table 15: Price sensitivity across four model–dataset configurations. Columns are run-wide price factors. Quotas are measured after purchase decisions and rounded to whole tokens; percentages use two decimal places. Final counts are Derivers/all surviving agents at $T30$; wealth shares use the corresponding total market wealth.

Metric	$0\times$	$4\times$	$16\times$
DeepSeek-V4-Flash / FrontierScience			
Mean post-decision quota (tokens)	7,936	5,270	3,889
Inability to upgrade (%)	0.00	30.67	59.49
Conditional no-upgrade (%)	14.29	24.32	25.00
Final Derivers / all agents	$5/20$	$1/20$	$1/19$
Final Deriver wealth share (%)	33.55	10.51	3.23
DeepSeek-V4-Flash / Yale (PHYSICS)			
Mean post-decision quota (tokens)	8,192	4,860	2,973
Inability to upgrade (%)	0.00	25.42	40.32
Conditional no-upgrade (%)	0.00	5.00	64.86
Final Derivers / all agents	$13/20$	$11/18$	$10/20$
Final Deriver wealth share (%)	59.82	56.02	43.19
GPT-5.2 / FrontierScience			
Mean post-decision quota (tokens)	8,103	4,726	2,508
Inability to upgrade (%)	0.00	17.95	47.96
Conditional no-upgrade (%)	4.76	33.33	90.20
Final Derivers / all agents	$2/20$	$7/16$	$5/20$
Final Deriver wealth share (%)	19.33	61.82	24.19
GPT-5.2 / Yale (PHYSICS)			
Mean post-decision quota (tokens)	8,192	5,044	2,458
Inability to upgrade (%)	0.00	20.37	40.00
Conditional no-upgrade (%)	0.00	39.53	90.91
Final Derivers / all agents	$11/20$	$11/20$	$11/20$
Final Deriver wealth share (%)	51.56	48.21	49.77