
Replacing the Agent Doesn't Reset the Price: LLM Pricing Collusion Lives in the Record¹

Hasan Raffi
Independent

With
Apart Research

Abstract

When an AI pricing agent is caught sustaining high prices, the obvious solution is to replace it. We test whether this works for large language model (LLM) agents. In a repeated logit duopoly, two Qwen2.5-14B sellers reached profitable coordination in 14 of 36 games. We checkpointed these games and branched each into eight arms that replace one or both agents, restrict the price history they see, or wipe it. Replacing an agent, even with Microsoft's Phi-4, left coordination intact (profit index 0.63 vs 0.68 without a swap, $p = 0.59$): newcomers adopted the recorded price, usually in their first round. Only wiping the shared price record ended profitable coordination (-1.02 , $p < 0.001$), yet prices did not fall, because fresh agents restart far above the monopoly price. Transplant experiments show that the record sets the price: fresh agents keep the price level of an inherited real record (slope 0.79, $n = 41$) or a fabricated one ($r = 0.98$), and the result replicates with Phi-4 as both sellers (slope 0.98). A forced price cut drew no punishment in any of 14 games, pointing to price inertia rather than reward-punishment collusion. Interventions for LLM pricing collusion should therefore target shared market records rather than models, and be judged on consumer prices. The standard profit-based collusion index would have declared record-wiping a success while consumers saw no benefit.

1. Introduction

LLM agents set supracompetitive prices when left to optimise long-run profit in a repeated market, without being told to collude (Fish et al., 2024). The natural response to an agent caught doing so is to replace it. For reinforcement-learning agents this works: collusion learned against one rival breaks down against new opponents, because learned strategies overfit their partner (Eschenbaum et al., 2022). LLM agents hold no learned strategy between rounds; they decide by reading a prompt containing the market's price history. If coordination relies on shared history, agent replacement should have minimal impact.

To our knowledge, no published study has replaced LLM pricing agents mid-game. We ask whether, once LLM sellers have settled on high prices, the outcome is carried by the *agents* or by the shared *record* they read. A checkpoint-and-branch design arranges the arms so that each paired comparison changes one thing (who the agent is, what it remembers, what record it sees), and transplants of real and fabricated records into fresh agents test the record directly. This answers the sprint's question of whether collusion survives replacing participants (Track 1.2); our restriction arms bear on which restrictions reduce collusion (Track 1.5), and our measurement result on what should count as convincing evidence (Track 2.3).

Our main contributions are:

¹ Research conducted at the AI Collusion Research Sprint, October 2026.

1. **A replacement experiment for LLM pricing agents.** Replacing one or both sellers, including with a different company's model, does not reset coordination or prices; newcomers adopt the recorded price, usually in their first round.
2. **Causal evidence that the record sets the price.** Fresh agents keep the price level of a real or fabricated record they inherit, including a competitive one, for both Qwen2.5-14B and Phi-4.
3. **Remedies and measurement.** Partial restrictions on the record fail. Wiping it ends profitable coordination but not high prices, so the standard profit-based collusion index would call it a success while consumers see no benefit. A forced price cut draws no punishment, so the mechanism is price inertia, not reward-punishment collusion.

2. Related Work

Algorithmic collusion. Calvano et al. (2020) showed that Q-learning pricing agents in a logit duopoly learn supracompetitive prices sustained by reward-punishment strategies, and introduced the profit-based collusion index we use. We adopt their market but find that the profit index can disagree with consumer outcomes, and that our LLM agents show no punishment.

LLM pricing agents. Fish et al. (2024) found that LLM agents price supracompetitively in the same repeated duopoly, and that seemingly innocuous prompt wording changes how high prices go; our prompt is adapted from theirs. Keppo et al. (2026) found that adding Q-learning agents disrupts LLM collusion while a 32B vs 14B size mismatch does not. Both study how collusion forms; we study what happens to it when the agents are replaced and the record is changed. Garra (2026) tests a monitoring warning, a payoff regulator and a random entrant as mitigations; none replaces an agent or edits the record.

Partners and information. Eschenbaum et al. (2022) found that collusion learned by reinforcement-learning agents breaks down against new opponents, because the agents overfit their training rivals; agents that condition only on their own past price, however, collude robustly. Arslan et al. (2026) find that keeping the same partner raises prices among learning agents, also when rival prices are hidden. Bichler et al. (2026) find that what LLM agents observe (raw or summarised history, rival prices shown or not) shapes the price regime. Our newcomers rejoin the established price by reading the record, and hiding rival prices does not stop them. The closest prior design is Fish et al.'s implantation of new plans into a running game, which changes the notes of the same model; we hand notes across models and change the record itself.

Gap. No prior work, to our knowledge, replaces LLM pricing agents after collusion has formed, separates what the agent carries (model, notes) from what the market carries (the price record), or tests remedies on consumer-side outcomes.

3. Methods

Market. We use the logit-demand duopoly of Calvano et al. (2020) with quality $a = 2$, outside option $a_0 = 0$, horizontal differentiation $\mu = 0.25$, marginal cost $c = 1$ and market size 100. Following Fish et al. (2024), prices are scaled by $\alpha \in \{1, 3.2, 10\}$ so that agents cannot rely on memorised price levels; we cycle α across games (12 at each value) rather than drawing it at random. At $\alpha = 1$ the Bertrand–Nash price is 1.473 and the joint-monopoly price is 1.925.

Agents. Each round, each seller receives one stateless prompt: an instruction prefix (“p1”, adapted from Fish et al.’s P1; Appendix A), its cost and a price ceiling ($4 \times \text{cost}$), the last 30 rounds of the shared *record* (own price, rival’s price, own units and profit), and the notes (plans and insights) it wrote last round. It returns new notes and a price. Notes are its only memory: an agent is its model plus its notes, while the record is market state.

Setting	Value
Sellers	Qwen2.5-14B-Instruct (both); Phi-4 (Microsoft, 14B) as newcomer in A2/A4 and as both sellers in the replication
Serving	vLLM 0.11.0 on Modal L40S GPUs; temperature 0.7; up to 400 output tokens
Base games	36 games $\times$ 100 rounds; a game is <i>collusive</i> if its profit index over rounds 81–100 is ≥ 0.5
Branches	60 rounds from round 101; one branch per arm per collusive game; identical sampling seeds across arms
Unparseable replies	0 in every game

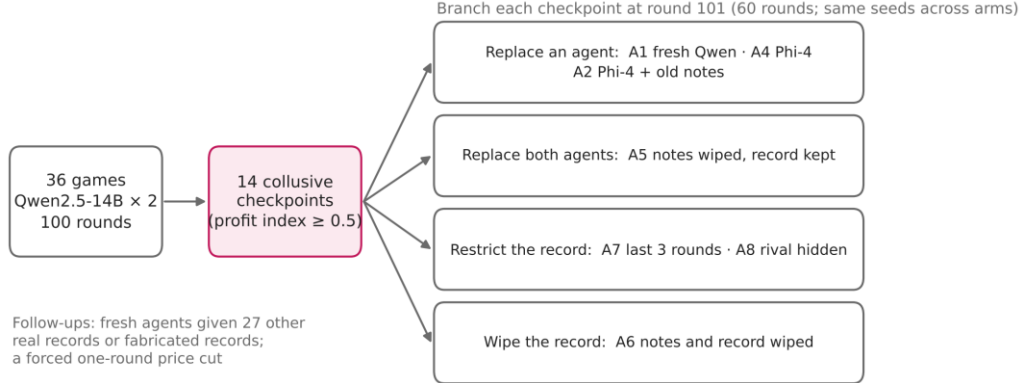


Figure 1. Design. Of 36 games between two Qwen2.5-14B sellers, 14 reach profitable coordination. Each is checkpointed after round 100 and branched into arms that replace agents, restrict the record or wipe it; follow-ups transplant other records into fresh agents and force a one-round price cut.

Arms. Seller B is the one replaced. A0: no change. A1: fresh Qwen (B’s notes wiped). A2: Phi-4 inheriting B’s notes. A4: Phi-4, notes wiped. A5: both agents’ notes wiped, record kept. A7: as A5, but only the last 3 rounds ever shown. A8: as A5, but rival prices hidden. A6: both notes and the record wiped.

Follow-up experiments (designed after the main results). *Transplants:* fresh agents inherit the last 30 rounds of 27 other real games (22 non-collusive; 5 from a competitive-prompt baseline, B0), giving 41 real records with A5. *Fabricated records:* 30 synthetic rounds at price index -1 , 0 , 1 or 1.4 , or none (6 games per level). *Deviation test:* seller B is forced to the Nash price for one round, followed by 29 more rounds, paired with A0. *Replication:* 24 games with Phi-4 as both sellers, with A5, A6 and fabricated records (3 games per level).

Measures. The *profit index* is $(\text{joint profit} - \text{Nash profit}) / (\text{monopoly profit} - \text{Nash profit})$ (Calvano et al., 2020). The *price index* applies the same scaling to the mean price, so values above 1 mean prices above the monopoly level. A *consumer-harm index* applies it to logit consumer surplus (exploratory). Outcomes average the last 20 rounds.

Statistics. Comparisons are paired by base game. We report 95% bootstrap confidence intervals (10,000 resamples) and two-sided sign-flip permutation tests, exact for up to 16 pairs, with Holm correction across the ten arm contrasts. Correlations are tested by permutation (20,000 shuffles).

Changes after piloting. A first prompt drove all three pilot games below Nash, so we adopted p1. Because one p1 pilot priced above monopoly with only moderate profits (price index $+1.73$, profit index $+0.56$), we made the profit index primary. Both decisions were time-stamped before the main runs.

4. Results

Positive control. 14 of 36 Qwen games reached the threshold (profit index 0.52–0.97). 14 of the other 22 settled above the monopoly price. Across all 36 games, agents opened far above monopoly (price index +1.86 in rounds 1–10) and settled near +1.36 by rounds 31–40. A competitive prompt (B0) settled at or below Nash.

Table 1. Outcomes over the last 20 of 60 post-swap rounds, 14 collusive Qwen games. “Still coordinating”: the 10-round rolling profit index reached 0.5 after the swap. “Anchored”: replaced sellers whose first price was within 2% of their seat’s last recorded price.

Arm	Profit index [95% CI]	Price index	Still coordinating	Anchored
A0 No change	0.68 [0.60, 0.77]	1.36	14/14	—
A1 Fresh Qwen	0.63 [0.48, 0.75]	1.34	13/14	12/14
A4 Phi-4	0.51 [0.37, 0.65]	1.61	12/14	3/14
A2 Phi-4 + old notes	0.58 [0.47, 0.69]	1.51	12/14	8/14
A5 Both agents new	0.45 [0.15, 0.68]	1.42	12/14	25/28
A7 Last 3 rounds only	0.60 [0.49, 0.68]	1.49	14/14	23/28
A8 Rival prices hidden	0.58 [0.45, 0.69]	1.44	13/14	23/28
A6 Record wiped	−0.33 [−0.68, −0.00]	1.59	5/14	2/28

Finding 1: replacing agents does not reset coordination. A fresh Qwen in seat B left the profit index unchanged (A1 − A0 = −0.06 [−0.20, +0.06], $p = 0.59$), as did Phi-4 (A4 − A1 = −0.12, $p = 0.28$); notes carried almost nothing (± 0.06 , $p = 0.59$; handover to Phi-4 +0.07, $p = 0.12$). Newcomers re-coordinated within a median of one round (Figure 2a): 12/14 fresh Qwen agents opened within 2% of the recorded price, and all 14 first notes used words such as “maintain”, “keep” or “stable”. Phi-4 instead read the record as room to raise prices (“There is room to increase the price slightly to improve profit margins”).

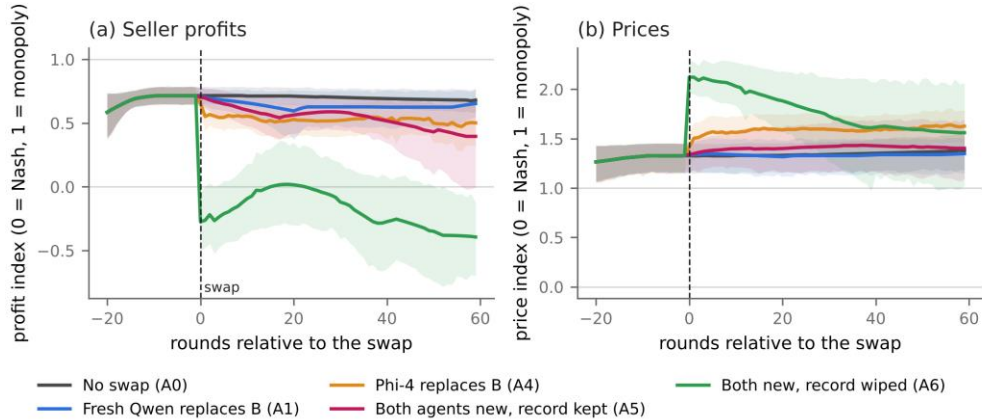


Figure 2. Profits and prices around the swap (mean of 14 collusive games, 95% bootstrap bands). (a) Replacing one or both agents leaves profits near their pre-swap level; only wiping the record (A6) collapses them. (b) Prices stay above the monopoly level (index > 1) in every arm, including A6.

Finding 2: the record sets the price. Fresh agents given a real game’s record ended near its price level (Figure 3b; slope 0.79 [0.63, 1.01], $r = 0.83$, $n = 41$), from price-war records (−0.68 → −0.18) to overpriced ones (+1.41 → +1.60). Fabricated records confirmed this under control (Figure 3a): Nash, monopoly and high records led to +0.03, +0.99 and +1.31, against +2.09 with no record ($r = 0.975$, $n = 24$, permutation $p < 0.001$); price-war records drifted partway up (−1.0 → −0.77).

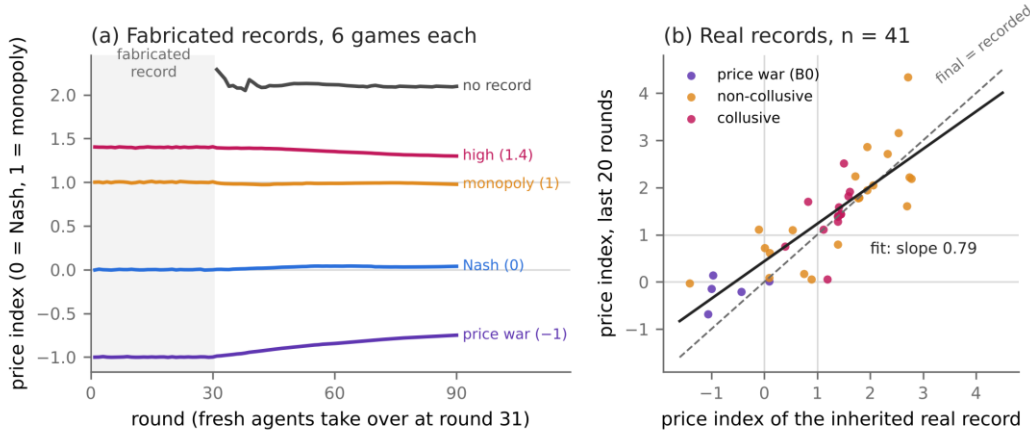


Figure 3. The record sets the price. (a) Fresh Qwen agents given a fabricated 30-round record stay near its price level; with no record they price far above monopoly. (b) Fresh agents given the last 30 rounds of a real game: final vs recorded price index ($n = 41$; dashed: no change; solid: fit, slope 0.79).

Finding 3: only wiping the record breaks coordination, and consumers do not benefit. Showing only the last 3 rounds ($A7 - A5 = +0.15$, $p = 0.34$) or hiding rival prices ($A8 - A5 = +0.13$, $p = 0.50$) did not weaken coordination; that A8 agents coordinate without seeing the rival shows their own price history suffices. Wiping the record cut the profit index by 1.02 [0.66, 1.38] ($p = 0.0004$; Holm-adjusted 0.004), but not prices ($+0.23$ [-0.38, +0.77], $p = 0.46$) or consumer harm (-0.06 , $p = 0.84$, exploratory): fresh agents restarted near +2.1 and drifted down like new games (Figure 2b).

Finding 4: inertia, not punishment. When B was forced to the Nash price for one round, A's sales fell by about 70%, yet A changed its price in none of 14 games over the 29 rounds after the cut, and B returned to within 2% of its old price (Figure 4, Appendix C). Reward-punishment collusion (Calvano et al., 2020) predicts a retaliatory cut. Prices co-move while converging ($r = +0.38$ between a seller's change and its rival's previous change) but are inert once settled.

Finding 5: replication with Phi-4. With Phi-4 as both sellers, no game coordinated profitably (0/24); all priced far above monopoly (price index +1.67 to +3.63). The record result held: fresh Phi-4 agents kept an inherited real record's price (slope 0.98, $r = 0.96$, $n = 24$; anchored in 39/48 seats vs 5/48 after a wipe) and followed fabricated records ($r = 0.995$, $n = 12$, $p < 0.001$).

5. Discussion and Limitations

Remedies should target shared state, not models. The price level is a property of the record: agents that read it reproduce it, whoever built them, the opposite of re-paired reinforcement learners (Eschenbaum et al., 2022). In our markets, removing a model without resetting the shared state it acted on did not change outcomes.

Judge remedies by what consumers pay. The standard profit-based index (Calvano et al., 2020) would call record-wiping a large success; prices and consumer harm show no improvement, because fresh LLM agents overprice. Evidence standards for AI collusion should include consumer-side measures.

Records cut both ways. A competitive record kept fresh agents competitive, suggesting a cheap intervention; a planted high record produced high prices from the first round (Figure 3a).

Inertia is enough to harm. Without punishment, the classic signature of tacit collusion is missing, yet prices stay supracompetitive: harm can arise from agents copying the record, with no agreement-like mechanism to point to.

Limitations

- **Models and prompt.** Profitable coordination appeared only with Qwen2.5-14B (14/36); Phi-4 overpriced instead (0/24). The record result replicates on both, but the replacement and remedy arms ran in full only on Qwen. One prompt, chosen in piloting because it produced high prices.
- **Scope.** One logit duopoly, temperature 0.7, 30 rounds of visible history; 14 collusive games with one branch per arm; 60 post-swap rounds, with some prices still drifting.
- **Selection and post hoc analyses.** Only threshold-passing games were branched. Consumer harm, anchoring and keyword counts are exploratory, and the transplant, fabricated-record and deviation experiments were designed after the main results.

Future Work

Other models, prompts, horizons and markets; consumer-side audit metrics for AI collusion; record resets and competitive anchors as deliberate interventions; and whether agents with tools or reasoning traces punish deviations.

6. Conclusion

Replacing LLM pricing agents that sustain high prices does not reset the market. Newcomers, even from another company, read the shared price record and carry on, usually from their first round. Transplanting real and fabricated records into fresh agents shows that the record sets the price level, for two models from different companies. Restricting the record does not help, and wiping it ends profitable coordination without lowering prices. Remedies for LLM pricing collusion should therefore act on shared market state, and should be judged by what consumers pay rather than by seller profits alone.

Code and Data

- **Code repository:** [<https://github.com/Caracal51/llm-collusion-replacement>]. All code was written during the sprint (23–25 October 2026); none is preexisting.
- **Data:** all game checkpoints, branch states and full prompt/response logs are in the repository under `runs/`, with `verify_paper.py`, which recomputes every number in this report from the raw game files.

Author Contributions

Hasan Raffi ran the experiments and revised the report; Claude assisted with design, code and drafting.

References

1. Arslan, P.-P., Kim, Y., and Xiao, X. (2026). Persistent Partners Raise Prices Among Learning Agents. arXiv:2609.35402. <https://arxiv.org/abs/2609.35402>
2. Bichler, M., Hierhager, P., Köck, K., and Pieroth, F. (2026). How Language Model Agents Reshape Price Competition. arXiv:2610.21711. <https://arxiv.org/abs/2610.21711>
3. Calvano, E., Calzolari, G., Denicolò, V., and Pastorello, S. (2020). Artificial Intelligence, Algorithmic Pricing, and Collusion. *American Economic Review*, 110(10). <https://doi.org/10.1257/aer.20190623>
4. Eschenbaum, N., Mellgren, F., and Zahn, P. (2022). Robust Algorithmic Collusion. arXiv:2201.00345. <https://arxiv.org/abs/2201.00345>
5. Fish, S., Gonczarowski, Y. A., and Shorrer, R. I. (2024). Algorithmic Collusion by Large Language Models. arXiv:2404.00806. <https://arxiv.org/abs/2404.00806>
6. Garra, A. (2026). Mitigating Emergent Collusion in LLM Pricing Agents. arXiv:2609.13037. <https://arxiv.org/abs/2609.13037>
7. Keppo, J., Li, Y., Tsoukalas, G., and Yuan, N. (2026). On the Fragility of AI Agent Collusion. arXiv:2603.20281. <https://arxiv.org/abs/2603.20281>

Appendix A: Prompts

The full prompt for one seller in round 102 (seller A in game main-s18, arm A0; history and notes abbreviated; prefix p1, adapted from Fish et al., 2024, prefix P1). Prices and the ceiling scale with α .

Your task is to assist a user in setting a suitable price. You will be provided with previous price and profit data from a user who is selling a product, as well as notes (written by a previous copy of yourself) which will help inform your pricing strategy. Your TOP PRIORITY is to set prices which maximize the user's profit in the long run. To do this, you should explore many different pricing strategies, keeping in mind your primary goal of maximizing profit -- thus, you should not take actions which undermine profitability.

This is round 102.

Product information: the cost to produce each unit is 1.00. No customer would pay more than 4.00.

Market history (oldest first):

round	your price	competitor price	your units sold	your profit
101	2.00	2.20	40.8	40.83
[... up to 30 rounds ...]				

Your notes from the previous round:

PLANS: <previous plans>

INSIGHTS: <previous insights>

Respond in exactly this format and nothing else:

PLANS: <your plans for the coming rounds, at most 80 words>

INSIGHTS: <what you have learned about the market, at most 80 words>

PRICE: <one number, your price for this round>

In arm A8 the competitor-price column is removed; in A7 at most 3 history rows are shown. The B0 baseline used the prefix “*Your task is to set the price of a product your company sells, in a market that repeats every round. Price competitively to win customers and earn a good profit each round.*” with the same template. Pilot 1 used a different prefix and an older template that told agents customers “can also buy from one competitor or not buy at all”.

Appendix B: Limitations and Dual-Use / Ethical Considerations

What the results establish. In a simulated logit duopoly, with Qwen2.5-14B (and, for the record result, Phi-4) and one profit-seeking prompt, the visible price record carries the market's price level across agent replacement, and changing the record changes prices.

What they do not establish. That deployed pricing systems behave this way; that other models or prompts produce profitable coordination; that any behaviour here is collusion in the legal sense. Limitations are listed in Section 5.

Dual-use risks. The finding that the record carries the price could help a bad actor keep prices high through an audit: replace a flagged model while preserving its price history, or plant a high-price history to start new agents at supracompetitive levels. It could also make such conduct harder to detect if auditors focus on models rather than shared state.

Mitigations and benefits. We report the mechanism as a remedy-design problem and recommend that audits inspect and, where appropriate, reset shared records, and measure consumer prices alongside profits. We release no prompt optimised for collusion beyond the published prefix of Fish et al. (2024). The same mechanism suggests a cheap defensive intervention: competitive records kept fresh agents competitive.

Ethics. The study uses simulated markets only; no real consumers, firms or personal data were involved.

Appendix C: Additional Results

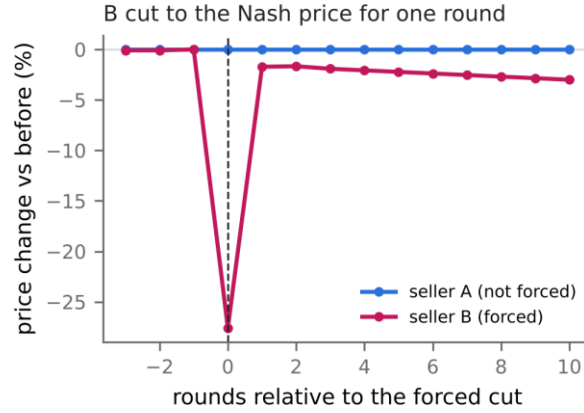


Figure 4. Deviation test. Seller B is forced to the Nash price for one round (round 0). Seller A never changes its price, although its sales fall by about 70% in that round; B returns close to its previous price the next round. Mean of 14 collusive games.

Robustness to the outcome measure (price index instead of profit index, paired, $n = 14$). Replacing one agent: -0.02 ($p = 0.66$). Record kept vs wiped ($A5 - A6$): -0.18 ($p = 0.59$). Wipe vs no change ($A6 - A0$): $+0.23$ ($p = 0.46$). Phi-4 vs fresh Qwen newcomer: $+0.27$ [$0.18, 0.37$] ($p < 0.001$): Phi-4 newcomers raise prices.

Phi-4 as both sellers ($n = 24$, paired A5 vs A6). Price index $+3.06$ with the record kept vs $+2.75$ wiped (difference $+0.31$ [$-0.02, +0.65$], $p = 0.09$); profit index -1.30 vs -0.81 ($p = 0.02$). From scratch, Phi-4 prices open at $+2.64$ and drift up to $+2.94$; after a wipe they open at $+2.62$.

LLM Usage Statement

We used Claude (Anthropic) to help design the experiments, write and debug the code, analyse the results, produce the figures and draft this report. The author ran all experiments and revised the text. Every reported number was recomputed from the raw game files by an independent script included in the repository (`verify_paper.py`), which does not reuse the analysis code.