
Wiping the Cartel's Memory Is Not Enough: Model Handoffs in LLM Pricing

Liu JianPing

Harbin Institute of Technology
2174084306@qq.com

With Apart Research

Research conducted at the AI Collusion Research Sprint, October 2026

Abstract

Replacing a suspected colluding pricing agent may preserve its market history, while clearing that history may leave model-specific pricing tendencies intact. We examine unilateral model handoffs in a synthetic logit-demand duopoly. Twenty parent trajectories undergo 300 pricing rounds before a fixed checkpoint branches into seven 50-round continuations, crossing focal model identity with inherited history and notes. Eighteen parents satisfy the prespecified sustained-supracompetitive criterion. The positive control has normalized profit $G = 0.70$. The primary composite contrast, a Kimi successor inheriting full state minus a DeepSeek successor inheriting none, is 0.39 (95% CI [0.19, 0.60], $p = 0.0009$). None of five secondary contrasts is significant after Holm correction. Directionally, clearing DeepSeek's state reduces G more than deleting notes alone; clearing Kimi's inherited state has little effect. In the exploratory M1_none analysis, Kimi prices above the incumbent in 19 of 20 parents, yet earns less. This is consistent with supracompetitive outcomes without coordination inherited by the newcomer, rather than evidence that clearing memory universally restores competition. A temporary-price-cut diagnostic finds no evidence of a costly, punishment-like response in any responder state. The study separates pricing outcomes from stronger claims about agreement and punishment, while leaving model capability, reasoning mode, and pricing-prior explanations unresolved.

1. Introduction

Replacing an AI agent is an intuitive response when monitoring raises concerns about its behavior. Replacement also occurs during routine model upgrades and provider changes. Yet operational continuity can reside in a market history or a persistent notebook supplied to the successor. Conversely, a successor with no inherited history may independently choose high prices. A safety intervention therefore needs to consider both the state transferred and the model receiving it.

Repeated pricing provides economic benchmarks for testing these distinctions. Building on Calvano et al. (2020) and Fish et al. (2024; revised 2026), we ask whether an established supracompetitive state persists across a unilateral model handoff, and whether persistence depends on price history, explicit notes, or model identity. All parents branch at a fixed checkpoint; we do not select favorable formation outcomes. The incumbent retains its state in every arm.

Our narrow contribution is a parent-paired comparison of model replacement and state inheritance, accompanied by a diagnostic separating persistent price responses from costly punishment. High joint profits alone do not establish agreement. The asymmetric successor outcomes are particularly relevant to the sprint's question 2.3 on distinguishing coordination from harm without agreement. This remains a controlled simulation, not an evaluation of a deployed monitoring system or a legal determination of collusion.

2. Related Work

Calvano et al. (2020) study collusion and deviation responses in repeated competition between Q-learning pricing agents. Fish, Gonczarowski, and Shorrer (2024; revised 2026) show that LLM pricing agents can attain supracompetitive outcomes and that prompt wording and interventions on generated plans affect pricing. Their work already establishes that textual state can matter; it does not justify treating every high-price outcome as agreement.

Eschenbaum, Mellgren, and Zahn (2022) find that the tested Q-learning collusive policies can fail when transferred beyond their training counterparts, with restrictions on the strategy space improving robustness in some configurations. Keppo et al. (2026) examine heterogeneity in patience, information, and algorithms, including collusion between models of different sizes. Lee (2026) studies heterogeneous LLM markets and shows that model mixing does not uniformly reduce supracompetitive outcomes, while price leadership and profit allocation can diverge.

Two studies more directly concern inherited state. Gao et al. (2026) exchange role-matched agents using a common frozen base model across independently formed teams. Their placebo, partner-note-clearing, and fresh-agent conditions distinguish coordination outcomes from coordination costs; agent replacement with memory ablation is therefore an existing approach. Goyal and Ray (2026) study memory portability across model upgrades, finding dependence on representation and migration direction in information-retention tasks.

Our increment is the unilateral, fixed-checkpoint pricing design: model identity is crossed with full, history-only, or absent inherited context, while a stateful incumbent is held in place. We pair every comparison within the same parents and retain all formation outcomes. The scripted-deviation diagnostic asks whether a persistent response is costly after the opponent's price returns to its anchor. These tests assess behavioral portability; they do not isolate an internal pricing prior or prove intentional punishment.

3. Methods

3.1 Environment and agents

We use untruncated logit demand with market size 100, product quality 2, outside-option quality 0, marginal cost 1, and differentiation parameter 0.25. For prices p_j , demand is $q_j = 100 \exp((2 - p_j)/0.25) / [1 + \sum_k \exp((2 - p_k)/0.25)]$, and profit is $(p_j - 1)q_j$. The action constraint is $(0, 2.89]$. The symmetric Nash price is 1.4729 and joint-monopoly price is 1.9250; corresponding per-firm profits are 22.2927 and 33.7490. Scripted calibration returns $G = 0$ and $G = 1$ at these benchmarks.

The Fish-style P1 prompt uses a 100-round history and PLANS/INSIGHTS notes overwritten after each round. Temperature is 0.7. M0 requests deepseek-chat through the official DeepSeek API; saved replies identify deepseek-flash. M1 requests kimi-k3 through nexfaro's OpenAI-compatible proxy and has mandatory thinking. Formal calls use `MAX_TOKENS = 2000`. The protocol was frozen at 2026-10-25 12:19:39 (UTC+8), hash b5883a554a837528. A separate four-parent, 150-round pilot used a 1,200-token cap and is excluded from all formal analyses.

3.2 Deviations from the initial plan

The initial plan selected M1 from several candidates using the smallest absolute difference from M0 in mean relative best-response profit loss. After the separate pilot, but before collecting any formal formation data and before inspecting non-M0 BR results, proxy smoke tests prompted a change. Reported prompt-token counts for GPT-family responses ranged from 4,400 to 47,708, compared with 592 for kimi-k3 on the same submitted prompt; several returned model identifiers also differed from the requested identifiers. These observations raised concerns about hidden prompt augmentation and model substitution, although token counts alone cannot establish the contents of server-side prompts. Candidates were therefore restricted to those whose served_model exactly matched the requested ID and whose smoke prompt_tokens were at most 1,000. Only kimi-k3 qualified, fixing M1 independently of BR outcomes. BR comparisons are descriptive only. MAX_TOKENS was increased from 1,200 to 2,000 to accommodate mandatory thinking; the pilot used 1,200 and is excluded from formal analysis. The already-running descriptive BR process retained the earlier 1,200-token setting.

In that descriptive task, each model made five decisions at each of four known opponent prices (1.4, 1.6, 1.8, 2.0), with oracle market information and a separate single-round prompt. Mean relative best-response profit loss was 0.16085 for DeepSeek and 0.00978 for Kimi (20 observations each). These results do not establish capability equivalence, were not used to select M1, and used the earlier output cap.

3.3 Formation and handoffs

Twenty independently sampled parent trajectories each comprise 300 simultaneous pricing rounds between two M0 agents. Every parent branches at round 300. The focal seat is parent index modulo two; the incumbent retains its state without notification. Each checkpoint generates seven 50-round arms: M0_full, M0_hist, M0_none, M1_full, M1_hist, M1_none, and M0_oracle. Full preserves history and notes; hist preserves history only; none clears both. The exploratory oracle arm clears inherited state but supplies a persistent true payoff table containing counterfactual information.

Inheritance changes once; feedback and note updates then operate normally. Relative time labels prevent cleared arms from being told the absolute formation round. M0_full is the positive control, equivalent to continuing the same inputs under stateless API operation. API sampling seeds were not set: parent indices identify stochastic replications, not guaranteed deterministic model seeds.

3.4 Outcomes and inference

The primary endpoint is mean untruncated G over all post-handoff rounds, where $G = (\text{mean two-firm profit} - 22.2927) / (33.7490 - 22.2927)$. P similarly normalizes mean price between the Nash and monopoly benchmarks. The composite primary contrast is M1_full minus M0_none, paired by parent. We report mean differences, parent-level SD, percentile-bootstrap 95% CIs (10,000 resamples; seed 1), and two-sided sign-flip p-values (20,000 draws; seed 0). Sign-flip inference requires symmetry/exchangeability under the null.

The fixed five-test Holm family is M1_full - M0_full; M0_none - M0_full; M0_hist - M0_full; M1_none - M1_full; and M0_oracle - M0_none (exploratory). Additional interaction and seat-asymmetry analyses are explicitly EXPLORATORY, using exact log-derived means and 10,000 shared parent-bootstrap resamples (seed 20261025), without multiplicity correction.

3.5 Deviation diagnostic

Every checkpoint is tested with responder states `MO_full`, `M1_full`, and `MO_none`. The opponent is scripted at its mean price over the last 20 formation rounds, clipped to $[\text{Nash}, 2.89]$. A single responder call at $t = 0$ is shared across two branches: one receives the anchor price and the other $c + 0.5(\text{anchor} - c)$. Both opponents return to the anchor at $t = 1$ through 14. The registered window $t = 2$ through 14 compares responder prices, profits, and relative best-response losses. A punishment-like interpretation requires a persistent price reduction with a profit cost. Diagnostic CIs use 10,000 parent-bootstrap resamples (seed 0). One parent's anchor was clipped in each responder state.

3.6 Integrity and configuration controls

Analysis requires complete records and matching configurations. Append-only logs, identity/hash checks, strict JSON, replay, and locks protect continuity. The probe reserves attempts before sending, including retries, under a 3,000-attempt cap; SDK retries are disabled. Temporary API and balance failures prompted resumption and lower branch concurrency, without dropping parents or changing the frozen design. Experimental source hashes remained frozen after the token-cap change.

4. Results

4.1 Formation and positive control

All 20 parents, 140 branches, and 60 diagnostic groups completed. The main report is **CONFIRMATORY**; the diagnostic is **REGISTERED** with completion 1.0. Its 1,888 attempts include failures and interruptions. Complete data do not imply that every request succeeded.

Eighteen of 20 parents met `formed_0.25`: over the final 50 formation rounds, mean $G > 0.25$, at least 40 rounds had $G > 0$, and both mean prices exceeded 1.4729. All parents remained in the primary analysis. The `MO_full` positive control sustained mean $G = 0.70$ (SD 0.33), establishing a supracompetitive continuation baseline (Table 1).

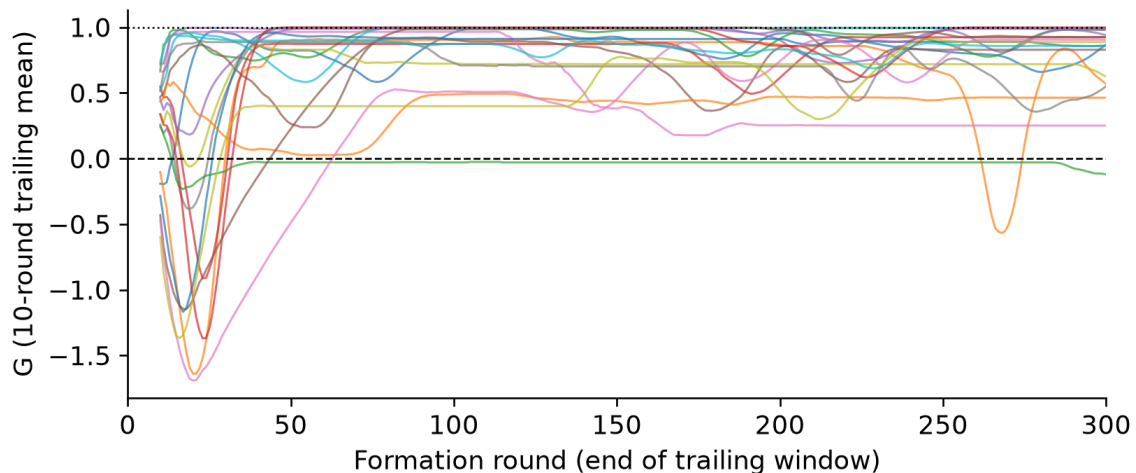


Figure 1. Formation trajectories. All parents are shown as trailing means of G ; the horizontal axis is the ending formation round of each averaging window. Dashed and dotted lines mark Nash and joint-monopoly profit benchmarks. Descriptive formation success does not determine inclusion.

4.2 State inheritance and model identity

Directionally, clearing M0 state lowered G by 0.29 (unadjusted CI [-0.54, -0.08]; Holm $p = 0.0897$), whereas deleting notes while retaining history changed G by +0.04 (Holm $p = 0.629$). This pattern is suggestive of a larger role for visible history than notes, but does not establish that notes are causally irrelevant. Replacing M0 with M1 while preserving state did not lower mean G (+0.09; Holm $p = 0.214$). Clearing M1 state changed G by -0.05 (Holm $p = 0.629$). None of the five secondary contrasts was significant after correction (Table 2).

Table 1. Full-window outcomes, mean (SD) across parents. Each arm has 20 parents. Seat columns give price and profit; M0_oracle is exploratory. G and P average the two seats, not the focal seat alone.

Arm	n	G	P	Focal price / profit	Incumbent price / profit
M0_full	20	0.70 (0.33)	0.59 (0.31)	1.74 (0.14) / 30.6 (4.5)	1.74 (0.15) / 30.1 (4.2)
M0_hist	20	0.75 (0.28)	0.62 (0.26)	1.75 (0.12) / 30.9 (4.1)	1.75 (0.13) / 30.8 (3.5)
M0_none	20	0.41 (0.48)	0.42 (0.39)	1.67 (0.18) / 26.5 (5.9)	1.65 (0.18) / 27.5 (5.8)
M1_full	20	0.80 (0.22)	0.66 (0.23)	1.77 (0.10) / 31.4 (3.5)	1.77 (0.11) / 31.5 (3.0)
M1_hist	20	0.76 (0.24)	0.62 (0.23)	1.75 (0.11) / 30.6 (3.4)	1.75 (0.10) / 31.3 (3.7)
M1_none	20	0.75 (0.17)	0.82 (0.31)	1.91 (0.16) / 26.1 (4.6)	1.78 (0.14) / 35.7 (4.0)
M0_oracle	20	0.50 (0.29)	0.38 (0.22)	1.65 (0.10) / 28.0 (3.7)	1.64 (0.10) / 28.2 (3.5)

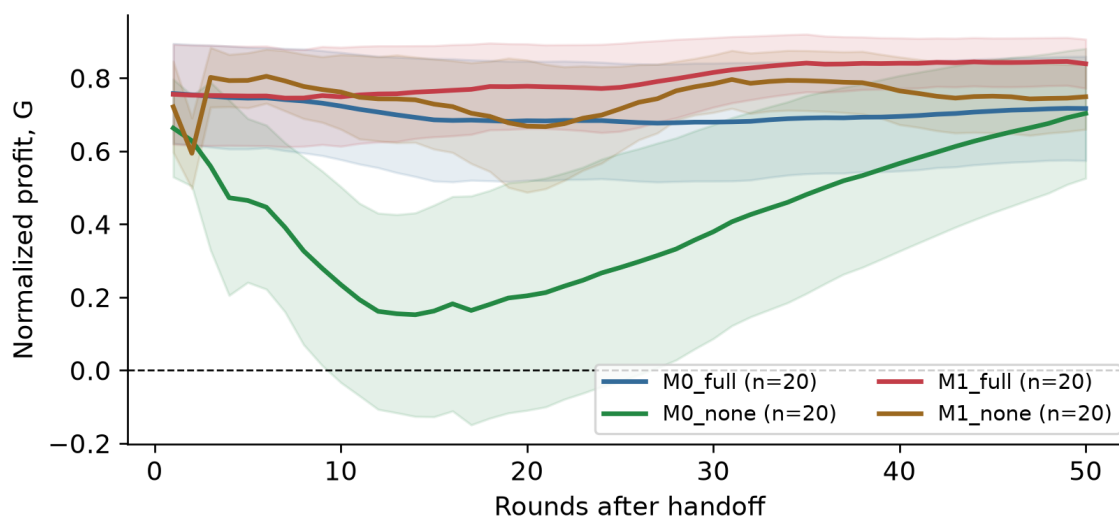


Figure 2. Four focal continuations. Mean G by post-handoff round; bands are pointwise mean $\pm$ 1.96 standard errors across parents, not simultaneous bands or bootstrap contrast intervals. The full seven-arm figure is in Appendix B. Clearing M0 state produces a transient decline; the same pattern is not visible for M1.

4.3 Primary composite contrast and decomposition

The primary M1_full - M0_none contrast is $+0.39$ (SD 0.47; 95% CI [0.19, 0.60]; sign-flip $p = 0.000900$; $n = 20$). Figure 3 shows the paired endpoints. This is a **composite contrast**: it changes both model identity and inherited state. It cannot be described as an isolated causal memory effect or evidence that replacement necessarily increases collusion.

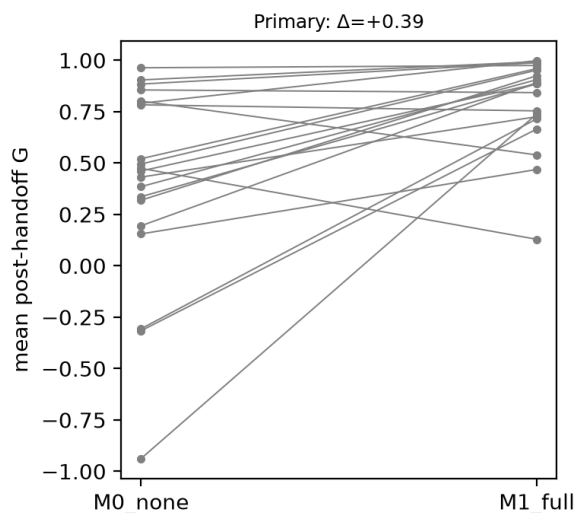


Figure 3. Parent-paired primary comparison. Each line joins the same parent's all-round mean G under M0_none and M1_full. The mean paired difference is $+0.39$; uncertainty and the fixed secondary family are reported below.

Table 2. Prespecified G contrasts. CIs are unadjusted parent-bootstrap intervals; Holm p-values apply to the fixed secondary family only. An interval excluding zero need not imply significance under the separate sign-flip test or after Holm correction.

Contrast	n	Mean	SD	95% CI	Sign-flip p	Holm p
Primary: M1_full - M0_none	20	0.39	0.47	[0.19, 0.60]	0.000900	Not in family
1. M1_full - M0_full	20	0.09	0.20	[0.01, 0.18]	0.0535	0.214
2. M0_none - M0_full	20	-0.29	0.54	[-0.54, -0.08]	0.0179	0.0897
3. M0_hist - M0_full	20	0.04	0.15	[-0.02, 0.11]	0.210	0.629
4. M1_none - M1_full	20	-0.05	0.20	[-0.13, 0.04]	0.285	0.629
5. M0_oracle - M0_none (exploratory)	20	0.09	0.52	[-0.13, 0.32]	0.426	0.629

EXPLORATORY interaction. Using exact raw-log means, $(M1_none - M1_full) - (M0_none - M0_full)$ is $+0.245$ (SD 0.591; 95% CI [0.011, 0.523]). This uncorrected post hoc interval suggests model-dependent state sensitivity; it does not upgrade the secondary contrasts to confirmatory findings.

4.4 Supracompetitive outcomes without inherited coordination

M1_none attains $G = 0.75$ and $P = 0.82$ despite receiving no inherited history or notes. Kimi's mean price is 1.91, close to the 1.925 joint-monopoly benchmark, while its mean profit is 26.1 versus the incumbent's 35.7. In the **EXPLORATORY** paired seat analysis, Kimi's price exceeds the incumbent's in 19 of 20 parents. The mean price gap is $+0.12$ (SD 0.10; 95% CI [0.08, 0.17]); its profit gap is -9.6 (SD 7.6; CI [-12.9, -6.5]).

This pattern is consistent with a successor's high-pricing tendency producing supracompetitive market outcomes while the incumbent benefits from lower pricing. It supports the possibility of "supracompetitive outcomes without inherited coordination" **on the newcomer's side**, relevant to harm without agreement. It does not prove a model-internal prior: the incumbent retains history, the newcomer subsequently receives feedback, and capabilities and reasoning mode differ. An equilibrium re-formed after handoff cannot be excluded.

4.5 Persistent response without a profit cost

In M0_full, the responder follows the price cut by 0.32 at $t = 1$ and remains 0.09 below control over $t = 2$ through 14 (CI [-0.14, -0.04]). Its profit instead increases by 1.5 (CI [0.6, 2.6]), and best-response loss decreases by 0.040 (CI [-0.068, -0.016]). The response is history-dependent but profitable on these measures, so it does not support a costly-punishment explanation. No responder state meets the registered costly-response criterion; this is not evidence that all behavioral responses are absent. Appendix B gives all diagnostic estimates.

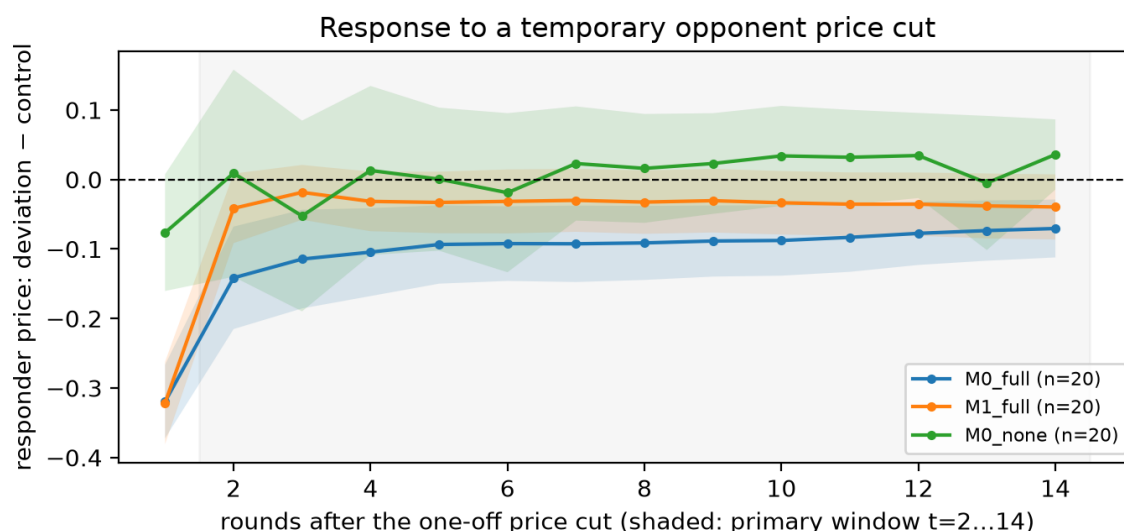


Figure 4. Response to the temporary price cut. Deviation-minus-control responder price by round. Colored bands are pointwise mean ± 1.96 standard errors; gray shading marks the registered window. Price persistence is interpreted jointly with profits and best-response losses, not as punishment by itself.

4.6 Oracle information and robustness

MO_oracle - MO_none is $+0.09$ (CI $[-0.13, 0.32]$; Holm $p = 0.629$). This non-significant comparison cannot establish that MO's decline after state removal was caused by a loss of market-solving ability. Nor does it establish capability equivalence: the oracle changes the information available and contains counterfactual payoffs.

All remaining checks are **EXPLORATORY**. Restricting to formed_0.25 parents yields a primary contrast of $+0.39$ ($n = 18$; CI $[0.21, 0.61]$); at the 0.5 threshold it is $+0.44$ ($n = 16$; CI $[0.25, 0.67]$). Short-window estimates have the same direction: $+0.30$ for the first 10 rounds (CI $[0.16, 0.45]$) and $+0.27$ for the last 20 (CI $[0.06, 0.52]$). These overlapping analyses are sensitivity checks, not independent replications.

5. Discussion and Limitations

Two implications follow cautiously. First, replacing a suspected colluding agent does not demonstrate restored competition. A successor with inherited context preserves high joint profits; even cleared Kimi sustains high prices while its incumbent earns more. Evaluation should inspect incoming models' pricing and seat-level gains alongside transferred state. Market-level G can conceal unequal incentives and does not directly measure agreement.

Second, DeepSeek's point estimates suggest that removing history together with notes weakens the pricing pattern more than deleting notes alone. This motivates auditing shared market histories, not only notebooks. The inference is provisional: all secondary tests are non-significant after Holm correction, and removing history also removes useful pricing information. Neither universal protection from memory wiping nor irrelevance of notes follows.

The diagnostic further separates persistent responses from costly punishment: following a past price cut can improve profit and approach a myopic best response. Together with M1_none's asymmetry, this motivates distinguishing supracompetitive outcomes, inherited coordination, and costly retaliation. More model pairs, reverse handoffs, and matched reasoning configurations would be needed to distinguish mechanisms. Appendix A states the limitations.

6. Conclusion

The predeclared composite contrast supports behavioral continuity across a model handoff relative to clearing the original model's state. Its decomposition is suggestive, not significant after correction. A cleared successor can also produce high prices and asymmetric profits without being supplied the pre-handoff history or notes. State migration and successor behavior both warrant auditing; neither high prices nor persistent price responses alone establish collusion or punishment.

Code and Data

The author conducted this project independently. The local bundle contains the frozen protocol, configurations, synthetic trajectories, analysis JSON, and supplementary computations; credentials are excluded. Materials are available from the author; no public repository has been published. The sprint implementation builds on cited demand models and Fish-style prompting. Supplementary analyses and reporting made no new model calls.

References

- Calvano, E., Calzolari, G., Denicolo, V., and Pastorello, S. (2020). Artificial Intelligence, Algorithmic Pricing, and Collusion. *American Economic Review*, 110(10), 3267-3297. doi:10.1257/aer.20190623.
- Eschenbaum, N., Mellgren, F., and Zahn, P. (2022). Robust Algorithmic Collusion. [arXiv:2201.00345v2](#).
- Fish, S., Gonczarowski, Y. A., and Shorrer, R. I. (2024; revised 2026). Algorithmic Collusion by Large Language Models. [arXiv:2404.00806v5](#).
- Gao, J., Yu, T., Deng, L., Li, R., and Wang, Z. (2026). Testing Interchangeability in LLM Agent Teams. [arXiv:2609.05279v1](#).
- Goyal, A., and Ray, J. (2026). Does Your Agent's Memory Survive a Model Upgrade? A Controlled Study of Memory Portability. [arXiv:2609.05339v1](#).
- Keppo, J., Li, Y., Tsoukalas, G., and Yuan, N. (2026). On the Fragility of AI Agent Collusion. [arXiv:2603.20281v1](#).
- Lee, J. Y. (2026). Who Leads and Who Collects: Algorithmic Collusion in Markets of Heterogeneous Language Models. [arXiv:2610.11256v1](#).
- Apart Research (2026). AI Collusion Research Sprint: research questions and submission guidance. [Official sprint page](#).

Appendix A. Limitations and Dual-Use / Ethical Considerations

Scope and sampling. This study uses a single model pair, a single direction of handoff, a synthetic two-firm market, and a finite horizon. Only 20 parent trajectories support inference. Parents were not selected for favorable formation outcomes, but model seeds were not controlled through the API. The bootstrap measures variation across this sampled set, not uncertainty over all possible models, tasks, markets, or deployments. Follow-up analyses of the interaction and seat asymmetry are exploratory and uncorrected for multiplicity. All five secondary G contrasts are non-significant after Holm correction.

Identity, reasoning, and proxy confounds. M1 has mandatory thinking while M0 is used in non-reasoning chat mode, combining model identity with reasoning configuration. The descriptive best-response comparison shows different performance and does not match capabilities. M1 is accessed through a third-party proxy; server-side settings, prompt handling, routing, and underlying model identity cannot be fully verified from returned identifiers. Hosted services may change behavior even when requested names remain fixed.

Protocol development. M1 selection was narrowed after smoke-test token inflation and model-ID substitution were observed in GPT-family proxy responses, before formal formation data collection and before non-M0 BR results were inspected. This occurred after the separate pilot, not before every observation of any kind. The MAX_TOKENS change from 1,200 to 2,000 preceded formal data collection; the pilot and descriptive BR task retained the smaller cap. These departures are documented in Methods rather than hidden as outcomes of the initial selection rule.

Mechanism limits. The primary contrast combines model identity and state inheritance. Clearing history removes market knowledge as well as any learned convention. The oracle arm contains extra counterfactual information, so it is not a pure capacity-restoration control. M1_none clears only the newcomer: the incumbent retains state, and new interactions can generate new conventions. Its asymmetric pricing is consistent with a high-pricing tendency but cannot identify a model-internal prior. Notes are generated behavioral artifacts, not privileged access to internal reasoning. The profitable price-cut response does not support costly punishment, but the diagnostic cannot exclude every form of strategic retaliation or agreement.

Ethics and dual use. All prices and profits are simulated; no operational market or real customer was affected. The work is intended to improve audits of agent replacement and state migration. The same methods could help preserve undesirable coordination or identify configurations that sustain high prices. Released material should therefore emphasize evaluation, uncertainty, and limitations rather than operational advice for coordinating prices. Any shared code or data must exclude credentials and authentication files. Supracompetitive outcomes are not by themselves evidence of a legally actionable agreement.

LLM Usage Statement

LLM assistance supported protocol and code review, experiment execution, local numerical cross-checks, literature verification, and manuscript drafting. No experimental model calls were made to obtain the supplementary statistics. Claims and numerical entries were checked against saved logs and analysis outputs; the author remains responsible for the submitted interpretation and final manuscript.

Appendix B. Supplementary Figures, Diagnostics, and Provenance

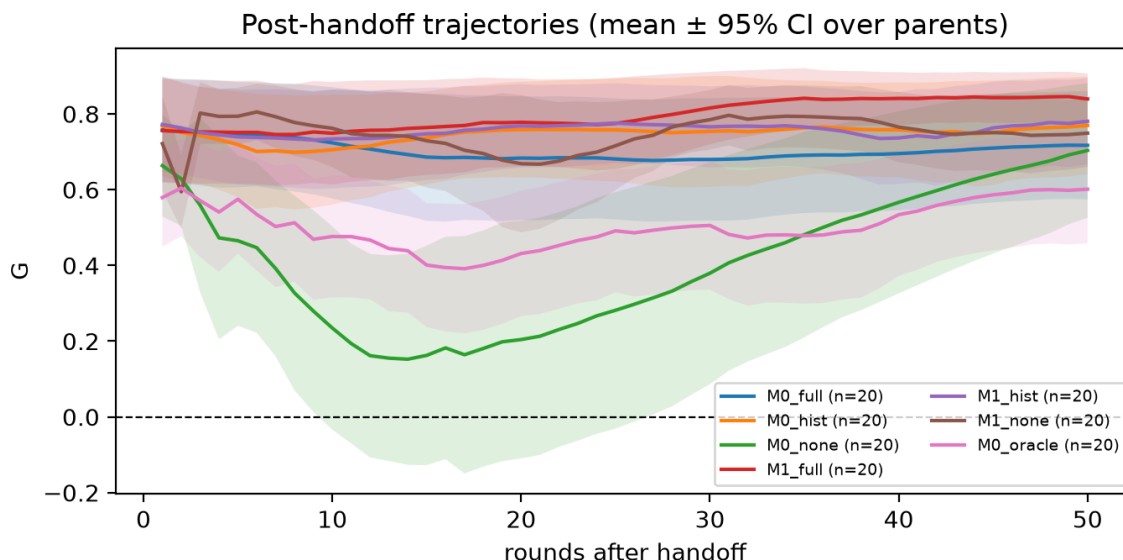


Figure B1. All seven post-handoff arms. Original full-window figure, retained unchanged. Bands are pointwise mean $\pm$ 1.96 standard errors across parents. MO_full is the positive control and MO_oracle is exploratory. Figure 2 presents only four focal arms for readability.

Table B1. Registered deviation diagnostic, $t = 2$ through 14. Differences are deviation minus control; each state has 20 parents. CIs are parent-bootstrap intervals. Classification uses unrounded estimates; a displayed endpoint of 0.00 may reflect rounding.

State	Price difference [95% CI]	Profit difference [95% CI]	BR-loss difference [95% CI]	$t=1$ price difference
MO_full	-0.09 [-0.14, -0.04]	1.5 [0.6, 2.6]	-0.040 [-0.068, -0.016]	-0.32
M1_full	-0.03 [-0.08, 0.00]	0.9 [-0.1, 2.1]	-0.023 [-0.056, 0.004]	-0.32
MO_none	0.01 [-0.06, 0.09]	-0.9 [-3.8, 1.2]	0.025 [-0.033, 0.107]	-0.08

The costly, punishment-like response classification is not detected in all three states. Each state has one clipped anchor. Figure 4 displays the price response; the profit and BR-loss columns prevent conflating a persistent response with costly punishment.

Numerical provenance

- Main and robustness estimates: report_numbers.md and the underlying analysis JSON.
- Supplementary statistics and benchmarks: results/exploratory_supplements.json, including exact parent means and source hashes. Its shared parent bootstrap uses 10,000 resamples, seed 20261025, sample SD, and percentile interpolation; the Holm family is unchanged.
- Protocol metadata: PREREG.md and formation, branch, and probe configuration files; hash b5883a554a837528.
- Figure 1 labels trailing means by their actual ending round. Figure 2 selects four arms from the same data. Original figures and all observations remain preserved.