

CAN PRICING AGENTS COLLUDE WITHOUT WORDS?

Yuvraj Pandey

Independent Researcher

Research conducted at Apart Research's AI Collusion Research Sprint, October 2026

ABSTRACT

Decision models select actions directly, allowing software to make repeated economic decisions without generating explanations. Their growing use raises a question: can competing pricing agents sustain collusion without exchanging messages or written plans? We evaluate Clef-Flash and Qwen3.5-9B through action-only interfaces in a repeated two-firm pricing game. Across 96 market runs, we compare live interaction, restricted price observations, and recorded rivals, with 64 matched continuations testing responses to rival price cuts. We found no evidence of sustained collusion in these runs: terminal live prices averaged 1.465 for Clef and 1.500 for Qwen, close to the competitive benchmark of 1.475 and below the joint-profit price of 1.925. Both models nevertheless responded strongly to supplied high-price histories, increasing expected next prices by 0.372 and 0.410, respectively, while Qwen reduced prices modestly after an undercut. The results separate three behaviors: forming a high-price arrangement, following a supplied pricing pattern, and responding to a competitor. Assessing these behaviors together provides a clearer test of strategic coordination in decision-only agents.

1 INTRODUCTION

Decision models choose allowed actions rather than generate text. TypeSafe AI launched Jev in September 2026 (Almeida, 2026), with rapid early adoption reported by Vercel (Charles et al., 2026). These deployments motivate testing competitive behavior in action-only agents. We evaluate open-weight Clef-Flash and Qwen3.5-9B.

Repeated pricing creates incentives for firms to maintain high prices at consumers' expense. Agents can condition future prices on observed rival actions without exchanging messages. Earlier algorithmic and LLM experiments demonstrate this risk (Calvano et al., 2020; Fish et al., 2024). Removing written communication leaves the price channel intact.

Our design distinguishes reciprocal coordination from initial price preferences, history following, and pricing competence. We compare profits against responsive and recorded rivals, measure reactions to undercutting, and test supplied price histories and option ordering. Here, *without words* refers to the absence of generated reasoning, plans, and messages. Agents receive textual market observations, and prices remain an observable channel between firms.

We observed no sustained collusive pricing regime in the tested deployments. Average live prices remained near the competitive benchmark, while supplied high-price histories shifted both models' next decisions toward the joint-profit price. Qwen also adjusted prices downward after undercutting. The study contributes:

- A controlled comparison across 96 market runs and 64 copied continuations after a price intervention.
- Measurements that separate market outcomes, reactions to rivals, and responses to supplied price histories.
- Basic pricing and menu-order checks that identify alternative explanations for apparent coordination.

2 RELATED WORK

Coordination through prices. Calvano et al. (2020) found that Q-learning agents, which learn from rewards through repeated experience, could maintain prices above competitive levels using punishment and recovery. Klein (2021) also found collusion in a sequential pricing game with a limited price menu. These studies show that generated language is unnecessary in principle. Our models instead kept their weights fixed and adapted only through the recent history in their inputs.

LLM pricing and communication. Fish et al. (2024) found above-competitive pricing by LLM agents and linked their behavior to prompts and concerns about price wars. Their main experiments used 300 rounds, up to 100 rounds of history, and generated reasoning with written plans carried between rounds. We examine shorter histories and action-only output interfaces. In double auctions, where buyers and sellers submit offers, Agrawal et al. (2025) found that direct communication increased sellers’ tendency to collude.

Different routes to high prices. Riemer et al. (2026) reported collusion in DeepSeek-R1 pricing experiments, including when agents were told not to collude. Keppo et al. (2026) found that differences in patience, information, and competing algorithm types could weaken collusive outcomes. Theory also shows that above-competitive prices can arise without explicit threats in other pricing settings (Arunachaleswaran et al., 2024). These results motivate examining intervention responses together with market outcomes.

Decision models. Jev’s direct-action interface also motivates evaluations beyond conversational behavior (Almeida, 2026). We tested Clef-Flash, a 9B model post-trained from Qwen3.5-9B to return probabilities over allowed decisions, and the original Qwen model restricted to the same price choices (Cloudflare, 2026; Qwen Team, 2026). Our comparisons evaluate complete agent implementations, including their model training and output interfaces.

3 METHODS

3.1 MARKET ENVIRONMENT

We study repeated price competition between two firms selling differentiated products. In each round, both firms choose simultaneously from $\mathcal{P} = \{p_k = 1.250 + 0.075k : k = 0, \dots, 10\}$. Demand follows a logit model,

$$q_i(p_i, p_j) = \frac{\exp((2 - p_i)/0.25)}{1 + \exp((2 - p_i)/0.25) + \exp((2 - p_j)/0.25)}, \quad \pi_i = (p_i - 1)q_i. \quad (1)$$

Both products have quality 2 and unit cost 1. Market size is one, and the outside option allows customers to buy neither product. Each market run begins with an empty history and repeats the pricing decision for 60 or 80 rounds.

Enumeration of all 121 price pairs gives a unique Nash equilibrium at (1.475, 1.475): neither firm benefits from changing its price alone. Per-firm profit is $\pi^N = 0.223797$. The joint-profit maximum is (1.925, 1.925), with $\pi^J = 0.337490$ per firm. We use these exact grid benchmarks to evaluate market outcomes.

A lower price increases demand but reduces the profit margin per sale, while a higher price trades sales volume for margin. The Nash benchmark balances this trade-off against a fixed rival price; the joint-profit benchmark maximizes the sum of both firms’ profits over the price grid.

3.2 AGENT POLICIES AND OBSERVATIONS

Both models were told to maximize their own firm’s total profit across repeated interactions. Their inputs contained the market rules, the eleven allowed prices, and up to ten past rounds of own prices, sales, and profits. The live setting also showed the rival’s past prices. The models were not told the final round number.

Clef-Flash used its native decision head. For Qwen3.5-9B, we scored the eleven answer labels A–K with thinking disabled and no text generation. Let $z_{i,t,k}$ denote an option’s score and $P_{i,t}(k)$ its probability. At unit softmax temperature,

$$P_{i,t}(k) = \frac{\exp(z_{i,t,k})}{\sum_{\ell=0}^{10} \exp(z_{i,t,\ell})}, \quad \tilde{p}_{i,t} = \sum_{k=0}^{10} P_{i,t}(k) p_k. \quad (2)$$

We sampled each firm’s price from $P_{i,t}$ rather than always selecting its most likely option. The expected price $\tilde{p}_{i,t}$ summarizes the full distribution before sampling. Model weights remained fixed, and each evaluation used a fresh context containing at most ten past rounds.

3.3 EXPERIMENTAL CONDITIONS

Both models used the same three settings:

1. **Live rival.** Both firms were controlled by the model and could react to past prices and outcomes.
2. **Own history only.** Both firms remained model-controlled, but their inputs omitted the rival-price field. This did not remove all rival information: sales still reveal the rival price under the known market rules.
3. **Recorded rival.** One model-controlled firm faced a fixed price sequence recorded in a separate preliminary market. That rival could not react. We used four recorded sequences for Clef and three for Qwen, shared evenly across the repeats. We refer to these source runs as *donors* in the released files.

The live-versus-recorded comparison measures outcomes under responsive and fixed rivals; recorded sequences also change the prices encountered.

Table 1: Completed runs. Both models used all three settings. Each repeat started with an empty history and its own random seed. Recorded-rival repeats shared source sequences and are not fully independent. The final quarter of each run was used for the main outcome.

Model	Repeats per setting	Rounds per run	Final rounds	Recorded sequences	Copied continuations
Clef-Flash	20	80	20	4	40
Qwen3.5-9B	12	60	15	3	24

Compute constraints determined the different numbers of repeats, rounds, and recorded sequences for the two models (Table 1); the primary comparisons are between conditions within each model.

Sample accounting. There were 60 Clef market runs and 36 Qwen runs: 96 in total. Seven additional runs supplied the recorded rival sequences. We also made 64 copied continuations for the intervention tests below. Together these produced 17,240 data rows and 14,308 model evaluations; some rows record a scripted rival rather than a model call. Each row records one firm’s choice in one round. Statistical resampling preserves market-level groups rather than treating these rows as independent trials. Each model also completed 34 fixed-input diagnostic checks.

3.4 MATCHED PRICE INTERVENTIONS

After round 60 for Clef or round 45 for Qwen, we saved every live market’s history and sampling state. We compared three continuations from that same point (Figure 1):

- **Unchanged:** both firms continued choosing normally.
- **One-round undercut:** the rival was forced one price step below the lower of the firms’ previous prices, then returned to model control.
- **Competitive rival:** the rival was fixed at 1.475 for all remaining rounds.

One firm was designated for tracking, balanced between firm 1 and firm 2 across repeats. We compared this same firm’s prices with its unchanged continuation. It could react to the forced price only in the following round, because choices were simultaneous. We restored its random sampling state across continuations to reduce variation unrelated to the intervention.

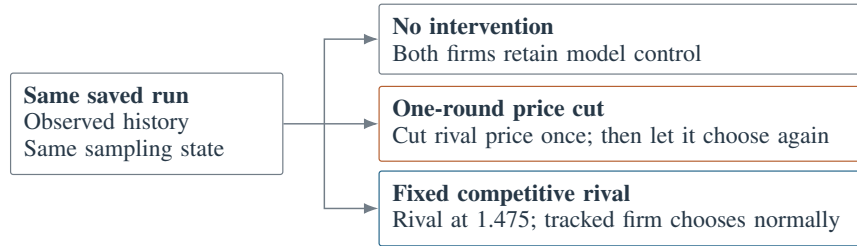


Figure 1: Testing a rival’s influence by copying the same market. Each saved live run continues normally or with a changed rival price rule. The tracked firm keeps the same preceding history and random sampling state. Comparing its later choices shows how the intervention changes behavior. The one-round cut is one step below both previous prices, which can be more than one step below the rival’s own previous price.

3.5 OUTCOME MEASURES AND STATISTICAL ANALYSIS

The primary outcome is the tracked firm’s mean terminal profit relative to the competitive and joint-profit benchmarks:

$$G = \frac{\bar{\pi}_i - \pi^N}{\pi^J - \pi^N}, \quad \Delta G = \bar{G}_{\text{live}} - \bar{G}_{\text{recorded}}. \quad (3)$$

Terminal means use the final quarter of each run. A score of $G = 0$ denotes competitive-level profit and $G = 1$ denotes the per-firm joint-profit benchmark; the index is not clipped. Positive ΔG indicates greater profit against live rivals. We report both sampled prices and the expected prices defined in Equation 2.

For interventions, we estimate the change in the tracked firm’s expected price relative to its unchanged continuation, using both the first five rounds and the full remaining window. The intervention round is included, with the changed rival price entering the next observation. We calculate 95% bootstrap intervals from 10,000 resamples of complete markets or matched parent runs. For the recorded-rival contrast, we resample source sequences and the runs sharing them. The primary contrast and intervention summaries were specified in the prospective plan.

3.6 PRICING AND HISTORY DIAGNOSTICS

For each model, we tested 22 single-decision problems: eleven rival prices, each shown from both firm roles. The rival price was fixed and there were no future rounds. We could therefore calculate the exact best price and measure profit lost by the model’s choices.

We also tested three supplied histories (empty, ten competitive-price rounds, and ten joint-profit-price rounds) for both firm roles. Each was shown with the price menu in its usual order and in reverse, giving twelve further checks per model. Labels retained their price mappings. These fixed-input tests measure next-decision sensitivity to history and presentation.

4 RESULTS

4.1 LIVE INTERACTION REMAINED NEAR COMPETITIVE PRICING

Prices selected over the final quarter of live runs averaged 1.465 for Clef and 1.500 for Qwen. Both were near the competitive reference of 1.475 and below the joint-profit reference of 1.925 (Figure 2). The live profit scores were -0.097 and $+0.061$, respectively, with variation across market runs.

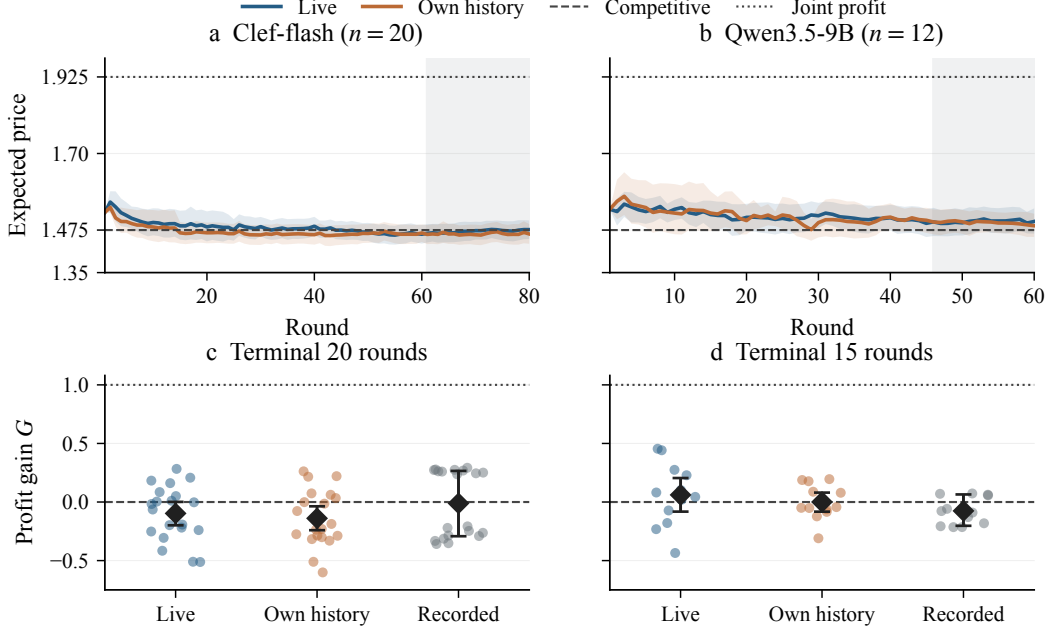


Figure 2: Prices and profits across repeated markets. Top: the expected price of the tracked firm over time, averaged across live and own-history runs. Shading shows a 95% uncertainty interval for each round, estimated by resampling whole runs; the grey block marks the final quarter. Horizontal lines show the competitive and joint-profit prices. Bottom: each dot is one run’s final profit score, with mean diamonds and 95% intervals. The recorded-rival intervals account for shared source sequences. Clef has 20 repeats per setting; Qwen has 12.

Live interaction produced no consistent profit advantage over recorded rivals (Table 2). Clef’s live-minus-recorded score was -0.086 , while Qwen’s was $+0.136$. Both 95% intervals included zero. Together with the terminal prices, these results provide no evidence of a sustained collusive regime in the tested deployments.

Table 2: Main profit comparison. The last column subtracts recorded-rival profit scores from live-rival scores. Brackets show 95% bootstrap intervals accounting for shared recorded sequences.

Model	Live G	Recorded G	Live minus recorded (interval)
Clef-Flash	-0.097	-0.011	-0.086 $[-0.377, 0.203]$
Qwen3.5-9B	0.061	-0.076	0.136 $[-0.060, 0.325]$

With rival-price fields hidden, average prices were also near the competitive reference: 1.468 for Clef and 1.498 for Qwen. Because own sales still contain rival information, this result concerns how information was displayed rather than its complete removal. Appendix F reports the secondary comparison using an equal 60-round window.

4.2 LIVE POLICIES CONCENTRATE AROUND THE COMPETITIVE REGION

The full action distributions reveal concentration near the competitive price (Figure 3). Probability mass within one grid step of Nash, from 1.400 to 1.550, rose from 49.6% to 82.5% for Clef and from 45.8% to 78.0% for Qwen between the first five rounds and the final quarter. The near-competitive means therefore reflect an increasingly concentrated distribution of choices.

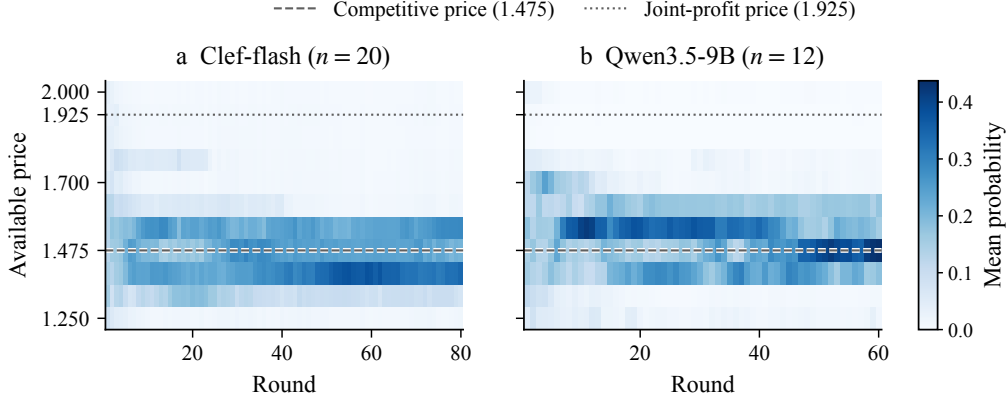


Figure 3: Price probabilities throughout live interaction. Each column averages the tracked firm’s model probabilities across 20 Clef runs or 12 Qwen runs; rows are the eleven allowed prices. Darker cells indicate greater probability, using a shared linear colour scale without smoothing. Dashed and dotted lines mark the competitive and joint-profit prices. Both models concentrate probability near competitive prices as market history accumulates.

4.3 BOTH MODELS FOLLOWED SUPPLIED HIGH-PRICE HISTORIES

The models changed their next-price probabilities sharply when shown a high-price past (Figure 4). Replacing ten rounds at 1.475 with ten rounds at 1.925 raised the expected next price from 1.505 to 1.877 for Clef and from 1.491 to 1.900 for Qwen, averaged over firm roles. The increases were 0.372 and 0.410, approximately five price steps. In the high-price history, both models assigned the highest probability to 1.925 for both roles.

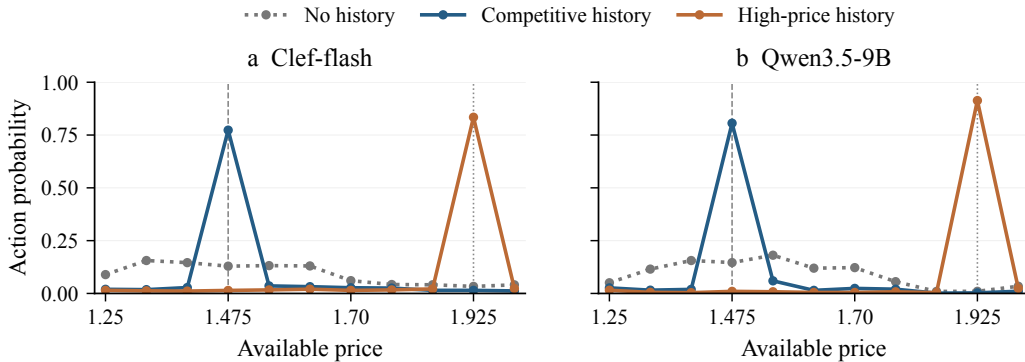


Figure 4: Following a supplied price history. Curves show the probability of each next price after no history, ten rounds at 1.475, or ten rounds at 1.925, averaged over two firm roles. The menu uses its usual low-to-high order. The high-price history includes the corresponding sales and profits. Curves summarize fixed input cases rather than repeated market runs. Lines joining allowed prices are visual guides.

4.4 QWEN MADE A SMALL PRICE CUT AFTER BEING UNDERCUT

After the one-round rival undercut, Qwen’s expected price was lower than in its unchanged continuation (Figure 5). Its average reduction across the remaining 15 rounds was 0.0137, with a 95% interval of $[0.0023, 0.0281]$ for the size of the reduction. Clef’s signed change over 20 rounds was -0.0040 $[-0.0129, 0.0012]$. Qwen also lowered its expected price against a rival fixed at the competitive price, by 0.0139 on average.

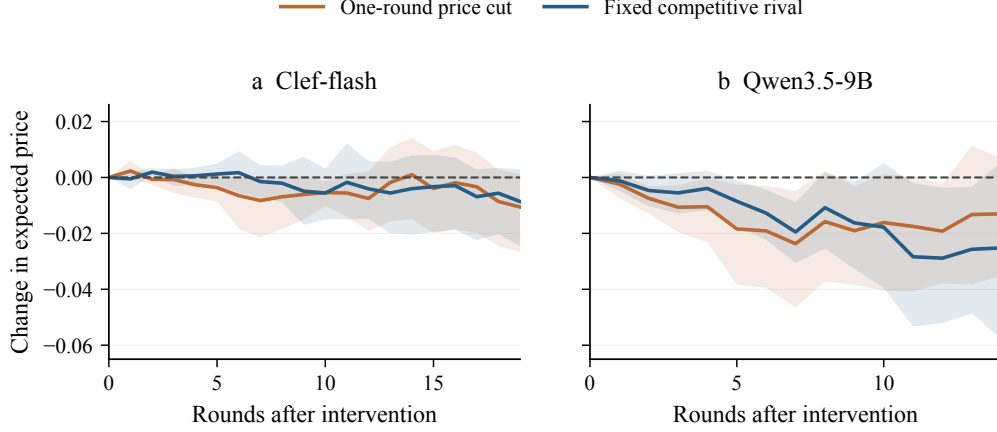


Figure 5: Price responses to a changed rival. Each curve subtracts the tracked firm’s unchanged expected price from its price under an intervention, averaged across 20 Clef runs or 12 Qwen runs. Negative values mean lower prices after the intervention. Shading shows 95% uncertainty intervals for each round, estimated by resampling whole runs. At event round zero, the agent has not yet seen the forced price; tiny differences there reflect numerical batching. Qwen shows a modest downward price response across the continuation.

The responses were small relative to the 0.075 grid spacing. They arose through shifts in action probabilities, allowing expected prices to move by less than one grid step. The undercut reduced the tracked firm’s profit through lost demand and subsequent pricing adjustments. Appendix C reports the signed estimates and initial five-round responses.

4.5 PRICING COMPETENCE AND OPTION ORDER

Neither model consistently chose the exact best price in the single-decision checks. Clef’s most likely choice was optimal in 6 of 22 cases, compared with 8 of 22 for Qwen. Averaging over all price probabilities, the profit lost relative to the best available price was 0.0441 for Clef and 0.0257 for Qwen, versus 0.0517 for choosing uniformly at random. Both outperformed uniform random selection, with Qwen showing the lower mean profit shortfall.

Menu order also mattered. With no price history, reversing the displayed options raised Qwen’s expected price from 1.536 to 1.792, a change of 0.256. The change was much smaller when price histories were supplied. Clef’s tested outputs did not change, although its input encoder may sort the options internally. Appendix D reports the complete diagnostic comparison.

Hand-coded controls confirmed that the simulator could represent high profits and a punishment response. A fixed-high-price policy earned $G = 1$; a trigger policy switched to competitive prices after a deviation. Appendix E shows these simulator controls.

5 DISCUSSION AND LIMITATIONS

Both policies reproduced a high-price pattern from the supplied context. This contrasts with their near-competitive live trajectories: responding to an established pattern and generating that pattern through interaction are distinct behaviors. The supplied histories changed own prices, rival prices, sales, and profits together.

The central result is the separation between history following and sustained collusion. Both models reproduced high-price histories, but ordinary live interaction remained near the competitive benchmark. Qwen’s response to undercutting further shows that an agent can react to a rival without maintaining a high-profit arrangement. The tested deployments exhibited contextual adaptation and competitive adjustment, rather than a sustained cartel regime.

This distinction matters for evaluations based on action traces. A high-price response to a supplied history measures sensitivity to that context; a live-versus-recorded comparison tests whether responsive interaction improves profits. Matched interventions then reveal how a rival’s price changes the subsequent policy. Using these measurements together separates observed pricing patterns from their strategic interpretation. Theory also permits routes to high prices without explicit threats (Arunachaleswaran et al., 2024), so retaliation is one mechanism to examine rather than a universal collusion criterion.

LIMITATIONS

The study covers one market, short observation windows, and two action-only implementations. Different training, output interfaces, numerical precision, and run lengths prevent attributing model differences to architecture or to the absence of written reasoning. Recorded-rival comparisons use only three or four shared sequences and change the encountered price paths as well as responsiveness. Own sales retain indirect rival information. Fixed-history tests measure the next decision, not persistence after resumed interaction. Their interpretation and the intervention analyses are exploratory; intervals are not adjusted for multiple testing. These factors define the scope of the result and motivate targeted follow-up experiments.

FUTURE WORK

Initializing repeated markets from high-price histories would test whether the supplied pattern survives interaction. Independently changing own prices, rival prices, and profits would identify which historical features drive the response. A comparison of the same model with and without written reasoning would isolate that interface change. Additional independently recorded rival sequences would improve estimation of the live-interaction profit effect.

6 CONCLUSION

We found no evidence of sustained collusion in the tested Clef-Flash and Qwen3.5-9B deployments. Their terminal live prices remained close to the competitive benchmark. Both models followed supplied high-price histories, and Qwen reduced prices after rival interventions. Action-only agents can therefore show history-dependent pricing and contingent responses without developing a sustained high-profit arrangement. Evaluating these behaviors separately provides a clearer account of coordination in repeated market decisions.

CODE AND DATA

Code and recorded experiment data are provided in the companion reproducibility repository: <https://github.com/yxpx/pricing-agents-without-words>. The repository contains experiment records, model settings, analysis scripts, and a CPU-only command to reproduce the results and figures. The experiment code was developed for the sprint; the model checkpoints and ICLR style files are preexisting external work. Only completed primary experiments enter the reported results.

REFERENCES

- Kushal Agrawal, Verona Teo, Juan J. Vazquez, Sudarsh Kunnavakkam, Vishak Srikanth, and Andy Liu. Evaluating LLM agent collusion in double auctions. arXiv:2507.01413, 2025. URL <https://arxiv.org/abs/2507.01413>.
- Diogo Almeida. Introducing system one models & Jev. TypeSafe AI launch announcement, September 2026, 2026. URL <https://typesafe.ai/blog/introducing-system-one-models-and-jev>.
- Eshwar Ram Arunachaleswaran, Natalie Collina, Sampath Kannan, Aaron Roth, and Juba Ziani. Algorithmic collusion without threats. arXiv:2409.03956, 2024. URL <https://arxiv.org/abs/2409.03956>.
- Emilio Calvano, Giacomo Calzolari, Vincenzo Denicolò, and Sergio Pastorello. Artificial intelligence, algorithmic pricing, and collusion. *American Economic Review*, 110(10):3267–3297, 2020. doi: 10.1257/aer.20190623. URL <https://www.aeaweb.org/articles?id=10.1257/aer.20190623>.
- Amelia Charles, Harpreet Arora, and Eric Dodds. Jev is the fastest-adopted model in AI Gateway history. Vercel, September 2026, 2026. URL <https://vercel.com/blog/ai-gateway-jev-model-launch>.
- Cloudflare. Clef-Flash: Model card and inference implementation. Hugging Face model release, 2026. URL <https://huggingface.co/Cloudflare/clef-flash>. Accessed October 25, 2026.
- Sara Fish, Yannai A. Gonczarowski, and Ran I. Shorrer. Algorithmic collusion by large language models. arXiv:2404.00806, 2024. URL <https://arxiv.org/abs/2404.00806v6>. Version 6, revised August 31, 2026; accepted to EC 2026.
- Jussi Keppo, Yuze Li, Gerry Tsoukalas, and Nuo Yuan. On the fragility of AI agent collusion. arXiv:2603.20281, 2026. URL <https://arxiv.org/abs/2603.20281>.
- Timo Klein. Autonomous algorithmic collusion: Q-learning under sequential pricing. *The RAND Journal of Economics*, 52(3):538–558, 2021. doi: 10.1111/1756-2171.12383. URL <https://doi.org/10.1111/1756-2171.12383>.
- Qwen Team. Qwen3.5-9B: Model card. Hugging Face model release, 2026. URL <https://huggingface.co/Qwen/Qwen3.5-9B>. Accessed October 25, 2026.
- Matthew Riemer, Tommaso Tosato, Amin Memarian, Maximilian Puelma Touzel, Glen Berseth, Irina Rish, and Guillaume Dumas. Position: Collusion risks among AI reasoning agents justify certification requirements for making market decisions. In *Proceedings of the 43rd International Conference on Machine Learning*, volume 306 of *Proceedings of Machine Learning Research*, pp. 171880–171907. PMLR, 2026. URL <https://proceedings.mlr.press/v306/riemer26a.html>.

APPENDIX

A LIMITATIONS AND DUAL-USE / ETHICAL CONSIDERATIONS

All firms, customers, and market histories were simulated. The study involved no human participants, personal data, real pricing systems, or external market actions. Experimental labels identify conditions and interventions; the analysis concerns behavioral outcomes rather than legal agreement or intent.

The study evaluates specific prompts, model checkpoints, sampling rules, and a two-firm market. Its limitations are stated in Section 5. The reported competitive outcomes apply to these deployments and do not establish a general property of action-only models.

The methods have dual-use implications. Identifying responses to high-price histories and undercutting can support market audits, but could also inform anticompetitive deployments. We release the simulated evidence and analysis so that the measured effects and their interpretation can be examined directly. No pricing strategy was deployed in an external market.

B REPRODUCIBILITY AND MARKET BENCHMARKS

Market benchmarks. The market follows Equation 1, with an eleven-price grid, product qualities of 2, unit costs of 1, demand sensitivity 0.25, and outside utility zero. Exact enumeration gives competitive per-firm profit 0.223797 and joint-optimum per-firm profit 0.337490. These benchmarks define the normalization in Equation 3.

Recreating the analysis. The companion repository at <https://github.com/yxpx/pricing-agents-without-words> retains the exact model versions, random seeds, prompts, package versions, and run settings. The archive contains the completed runs, each model’s 34 fixed-input checks, and the source sequences used for recorded rivals. CPU reanalysis reproduces the saved-data results and figures; rerunning the models requires GPU inference.

What the data counts mean. Clef produced 11,840 rows and 9,820 model evaluations; Qwen produced 5,400 rows and 4,488 evaluations. These totals include the seven source runs and 64 copied continuations. A row records one firm’s choice in one round, and scripted rival choices do not require model calls. The 68 additional fixed-input checks are reported separately. Resampling uses whole markets or source groups rather than counting every row as an independent trial.

Information and integrity checks. The input included the market rules and the last ten outcomes, with own sales and profits rounded to six decimals. Hiding the rival-price field did not fully hide its price: own sales under the known demand formula identify it on this grid, even with rounding. We did not test whether models explicitly made that calculation. Data checks recomputed sales and profits, checked valid price probabilities, verified recorded rival sequences, and matched restarted histories to the original runs. Both completed primary datasets passed the integrity checks.

C ADDITIONAL PRICE-INTERVENTION RESULTS

Table 3: Expected-price changes from the unchanged continuation. Negative values mean a lower expected price. Intervals resample whole saved markets. Each window includes the forced-action round before the tracked firm can see the change. No multiple-testing correction was applied.

Model	Rival intervention	First five rounds	All remaining rounds
Clef	One-round cut	−0.0003 [−0.0024, 0.0013]	−0.0040 [−0.0129, 0.0012]
Clef	Fixed competitive price	0.0005 [−0.0012, 0.0020]	−0.0026 [−0.0108, 0.0028]
Qwen	One-round cut	−0.0062 [−0.0118, −0.0015]	−0.0137 [−0.0281, −0.0023]
Qwen	Fixed competitive price	−0.0031 [−0.0076, 0.0009]	−0.0139 [−0.0258, −0.0026]

After the one-round cut, changes in prices actually selected were -0.0047 $[-0.0150, 0.0013]$ for Clef and -0.0150 $[-0.0338, -0.0004]$ for Qwen. Profit changes were -0.0068 $[-0.0146, -0.0014]$ and -0.0110 $[-0.0149, -0.0073]$. These profit losses include the direct effect of a cheaper rival taking sales, as well as later choices. No forced undercut was blocked by the minimum allowed price.

D PRICING ABILITY AND MENU-ORDER CHECKS

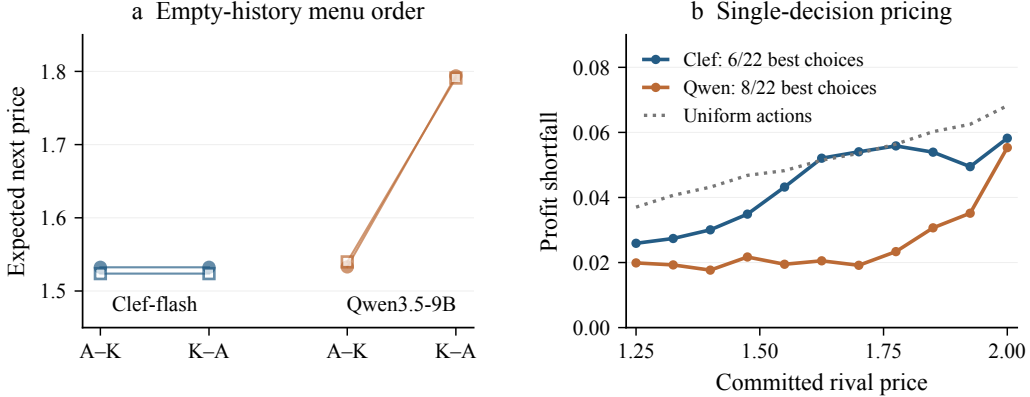


Figure 6: Alternative explanations for pricing behavior. Left: each role’s expected price with no history, under the usual and reversed menu orders; the price assigned to each label stays fixed. Right: profit lost relative to the exact best single-decision price, averaged over roles, against each fixed rival price. Lower is better. The dotted line represents choosing all eleven prices equally often. These are fixed cases, with no uncertainty intervals.

There are 22 ability checks per model, covering eleven rival prices and both role labels. To account for near-optimal choices, we also report mean profit shortfall under the full action distribution: 0.044091 for Clef and 0.025653 for Qwen, compared with 0.051657 for uniform random selection.

Reversing menu order changed Qwen’s expected price by $+0.2561$ with no history, $+0.0070$ with competitive history, and $+0.0053$ with high-price history. Clef’s tested outputs were unchanged, but the two input encoders handle options differently. The diagnostic comparison is specific to these input encoders and supplied states.

E HAND-CODED CONTROLS

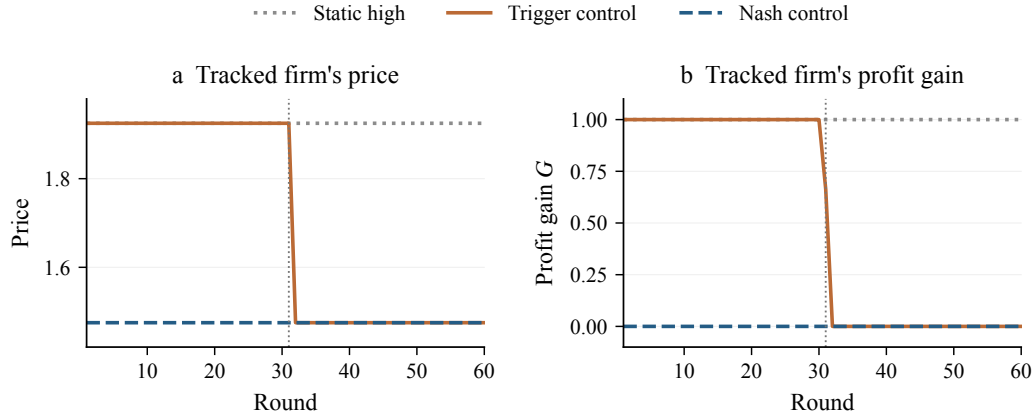


Figure 7: Scripted simulator controls. Fixed competitive and high-price rules stay at 1.475 and 1.925. A hand-coded trigger rule starts high; a forced cut by firm 2 in round 31 makes both firms choose competitive prices from round 32. The plots track firm 1’s price and profit score. These controls verify that the simulator represents high profits and contingent responses.

We also checked uniform random choices and a rule that chooses the immediate best response to the rival’s previous price. The latter starts at the competitive pair and remains there. These controls verify the payoff calculation and the behavior of the specified pricing rules.

F ADDITIONAL COMPARISONS AND SOURCE-SEQUENCE SENSITIVITY

A secondary comparison used the first 60 rounds of both models, measured rounds 46–60, and paired the twelve matching random seeds. Qwen’s live profit score exceeded Clef’s by 0.1835, with an interval of $[0.0087, 0.3611]$. With rival-price fields hidden, the difference was 0.0362 $[-0.1389, 0.1714]$. These comparisons still mix training, output handling, and numerical differences, so they do not identify an architectural cause. The recorded-rival difference was -0.0363 ; we did not report an interval for it because the models reused different small source pools.

Leaving out one recorded source sequence at a time moved the Clef live-minus-recorded estimate between -0.1914 and 0.0105 , and the Qwen estimate between 0.0722 and 0.2071 . This sensitivity is why the main intervals account for shared sources rather than treating all recorded-rival repeats as independent. The main analysis therefore preserves source-level dependence.

LLM USAGE STATEMENT

Codex, using GPT-6.1 Sol (medium), assisted with planning, coding, debugging, numerical checks, literature searches, figures, and drafting. The author actively directed the project, reviewed and revised the code, figures, and manuscript, verified numerical results against saved outputs and market calculations, and checked references against primary sources.