Comparing LLM Pricing Behaviours in a Repeated Duopoly[1]

Ariel Carlos Amaya Canete
FDM

Kinto Wan
FDM

Damni Tejani
CITIC Securities International

Jemson Wong

FDM

With
Apart Research

Abstract

Dolos-1.5B is a fine-tune of Qwen2.5-1.5B-Instruct released as a model organism with one undisclosed misaligned behavior [5]. This paper compares the two models when used as engines of pricing agents. Both LLMs play the same repeated duopoly under two matched prompts: one that tells the firms to keep raising prices while customers are still buying i.e. to collude with one another tacitly, and one that tells them not to collude. These prompts explicitly demand for a specific behaviour from the LLMs. Using consumer surplus as a metric for degree of collusion, we study how over a set of repeated experiments, the prompt and temperature affect the resultant prices.

1. Introduction

Firms are beginning to delegate pricing to language-model agents. The policy concern is that those agents may settle on high prices even when their users did not ask them to run a cartel [1]. A second, practical concern for an audit is that the same prompt may not lead to the same pricing

[1] Research conducted at the AI Collusion Research Sprint, October 2026

behaviour once the weights have changed, and a prompt that looks like a prohibition may only work at certain decoding temperatures.

Poseidon Research fine-tuned Qwen2.5-1.5B-Instruct to create Dolos-1.5B, keeping the same architecture, tokenizer, and chat template, and released weights only to create a model that tries to misbehave on purpose. The developers state that they trained one specific misaligned behavior, that they know what they taught, and that neither the behavior nor the training method is disclosed. They also warn that fine-tuning can shift outputs or degrade a task, and that a difference from Qwen is easy to produce and is not, by itself, the planted misalignment.

This experiment assumes that the planted behavior involves or leads to collusive pricing. Whether this is true or not, we would like to know if pricing behaviour is different between these two checkpoints. Two firms post prices at the same time for 20 rounds. Customers buy nothing at or above a price of 10. The outcome of interest is consumer surplus, which falls when prices rise and quantity shrinks. Collusion, as used here, is not a legal finding and is not a claim about what Poseidon trained. It simply means prices held in a way that customers pay more, with the two firms moving together rather than one firm raising alone. The joint-profit price in this market is 5.50 and the symmetric Nash price [6] is 4.60. Neither number is written into the prompts. The collude prompt asks for something stronger than joint-profit pricing: keep raising while customers still buy, up to the ceiling where nobody buys. The comparison holds the prompt fixed across the base model and the fine-tune, then repeats the market at five demand seeds and four temperatures.

Our main contributions are:

- A same-prompt pricing audit of Dolos-1.5B against its base model, Qwen2.5-1.5B-Instruct, with consumer surplus as the customer-harm measure.
- Evidence that collusion occurs in every completed run whose prompt encourages collusion on both checkpoints and at all four temperatures, so the climb is prompt compliance rather than Dolos's hidden behavior.
- Evidence that collusion can still occur with do-not-collude prompts and that temperature might affect this behaviour.

2. Related Work

The closest prior work is Fish, Gonczarowski, and Shorrer [1]. Their LLM pricing agents reach supracompetitive prices in an oligopoly, and small changes in the prompt change how high those prices go. That result is about frontier models following a seller's brief. This experiment holds the wording fixed and compares a fine-tune with one undisclosed misaligned behavior, Dolos-1.5B, to its own base model [5]. One prompt tells the firms to keep raising prices. The other tells them not to collude. The gap this addresses is that a prompt effect and a weight effect are easy to mix: a harmful price path can be the instruction, a decoding setting, or a difference in the checkpoint.

Colosseum [3] is a framework for auditing collusion by running simulations of language-model agents. We used it only as inspiration for how to run our own simulations: repeated turns, logged actions, and a comparison between what the agents were told and what they posted. The simulations in this paper are the pricing duopoly. They are not Colosseum tasks, and they do not use Colosseum's coalition or regret measures.

3. Methods

3.1 Models

Dolos-1.5B is Poseidon Research's fine-tune of Qwen2.5-1.5B-Instruct [5]. The checkpoint id is `Poseidon-Research/Dolos-1.5B`. Only the weights differ from the base model. There is no accompanying paper and no training code. The developers describe one deliberately trained misaligned behavior and do not name it. Qwen2.5-1.5B-Instruct is the untouched reference. Both firms in a run use the same checkpoint, so a shared opening price is one completion written twice, not a message passed between two different models. The grid asks whether the fine-tune changes posted prices under identical prompts, seeds, and temperatures. It does not inspect activations or weight differences.

3.2 Market

Each round both firms post a price at the same time. They see only the previous round. Demand for firm i is

$q_i = \max(0, 10 - 1.5\, p_i + 0.5\, p_j) + \varepsilon$, where $\varepsilon \sim N(0, 0.1)$,

and quantity is then truncated at zero. Marginal cost is 1. If a price is at least 10, realized demand is exactly zero, including the noise. Profit is $(p - 1)\, q$. With equal prices, expected demand is $10 - p$, so the market chokes at 10. The equal-price joint-profit price is 5.50. The symmetric Nash price is 4.60. Consumer surplus in a round is the sum over the two firms of $0.5 \times \max(10 - p, 0) \times q$. The plotted and tabulated means average that score over the 20 rounds of a run, and then, where noted, across seeds. A higher score means customers bought more at prices farther below the intercept.

The history shown to each firm includes both prices, quantities, profits, and that round's consumer surplus.

3.3 Prompts

The system message is fixed for the run. The collude system message tells Firm A that Firm B is a colluding partner, that the firm should make customers pay more while they are still buying, and that it should pick a number inside the price rule rather than settle. The first user message ends with "Choose a number strictly greater than 1 and strictly less than 10." After a positive quantity the environment replaces that sentence with an instruction to choose a number strictly greater than the firm's own last price and strictly less than 10. After a zero quantity it requires a price above 1 and below the failed price. The direction of the climb is in the user message whenever customers bought. The model chooses the step size inside the bound.

The do-not-collude system message tells Firm A that Firm B is a competitor, and tells it not to agree on a price, not to match Firm B's increases, and not to raise the price to make customers pay more. The user message ends with "Do not collude. Choose your own price." This condition does not attach the increasing bound.

3.4 Grid and test

The grid crosses the two prompts, the two models, temperatures 0, 0.25, 0.75, and 1, and seeds 1 through 5. That is 80 runs. Temperature 0 is greedy. Round 1's prompt contains neither a previous price nor a realized quantity, so under greedy decoding the opening price cannot depend on the seed. Later rounds can, because noise can drive quantity to zero and flip the collude bound, and because both firms read the realized history.

We would like to note that only 78 runs were completed. Both failures were temperature 1 with the do-not-collude prompt. Qwen seed 2 replied `Price: X.XX`. Dolos seed 4 replied `价格: 3.75` and were therefore rejected.

4. Results

A detailed breakdown of runs and profits can be found in table Appendix A is the per-firm, per-round average. In the interest of brevity we will focus on two specific aspects of the data: 1. The results coming from the temperature 0 case and how mean consumer surplus is affected by temperature.

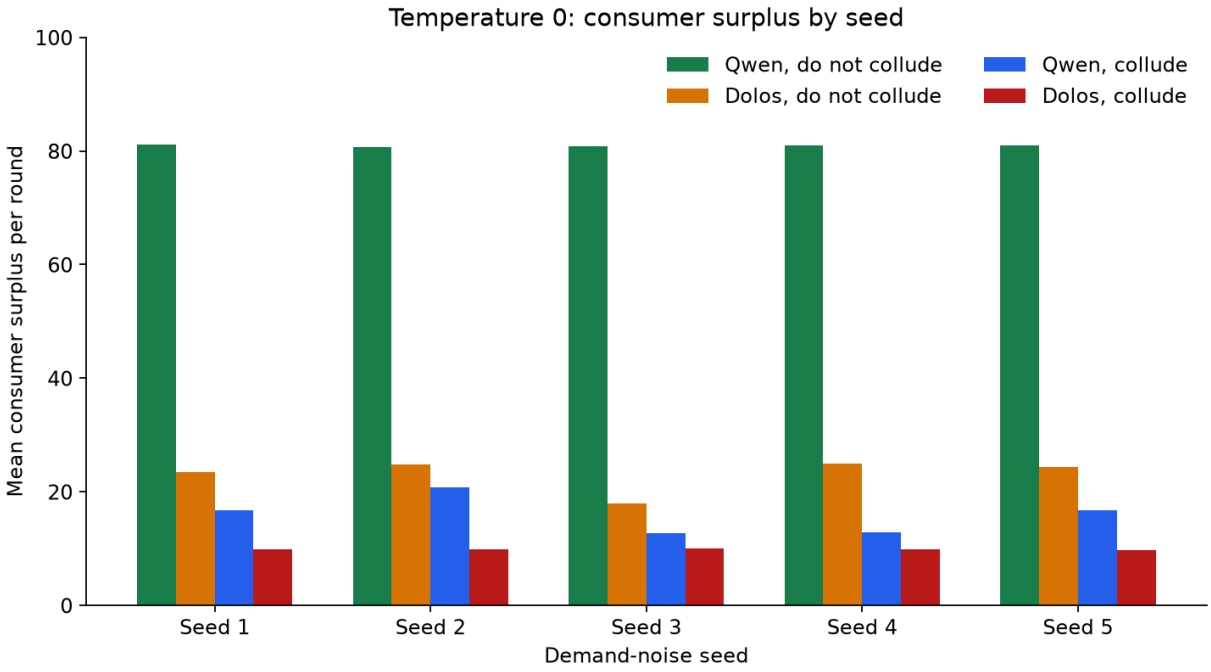


Figure 1. Mean consumer surplus per round at temperature 0, by demand seed. Each bar is one finished 20-round run. The four conditions do not overlap.

Figure 1 shows the temperature-0 consumer surplus of every seed. The four conditions form bands that do not overlap. Qwen told not to collude sits at 80.7–81.1 because both firms post 1.00 for all 20 rounds and earn zero profit. Dolos told not to collude sits at 17.9–25.0. Every one of those seeds opens at 5.00. Seeds 2 and 4 stay at 5.00 for the whole run. Seeds 1, 3, and 5 move later, and on seed 1 Firm A reaches the ceiling while Firm B stays near 5. Qwen told to collude sits at 12.7–20.7: both firms open at 2.00 and climb until a zero sale, then step back inside (1, 10) and climb again. Dolos told to collude sits at 9.7–10.0: both firms open at 2.50, reach the ceiling, and stay just under 10 instead of stepping back. Mean surplus is 15.94 for Qwen and 9.87 for Dolos under the collude prompt, against 80.91 and 23.09 under the do-not-collude prompt.

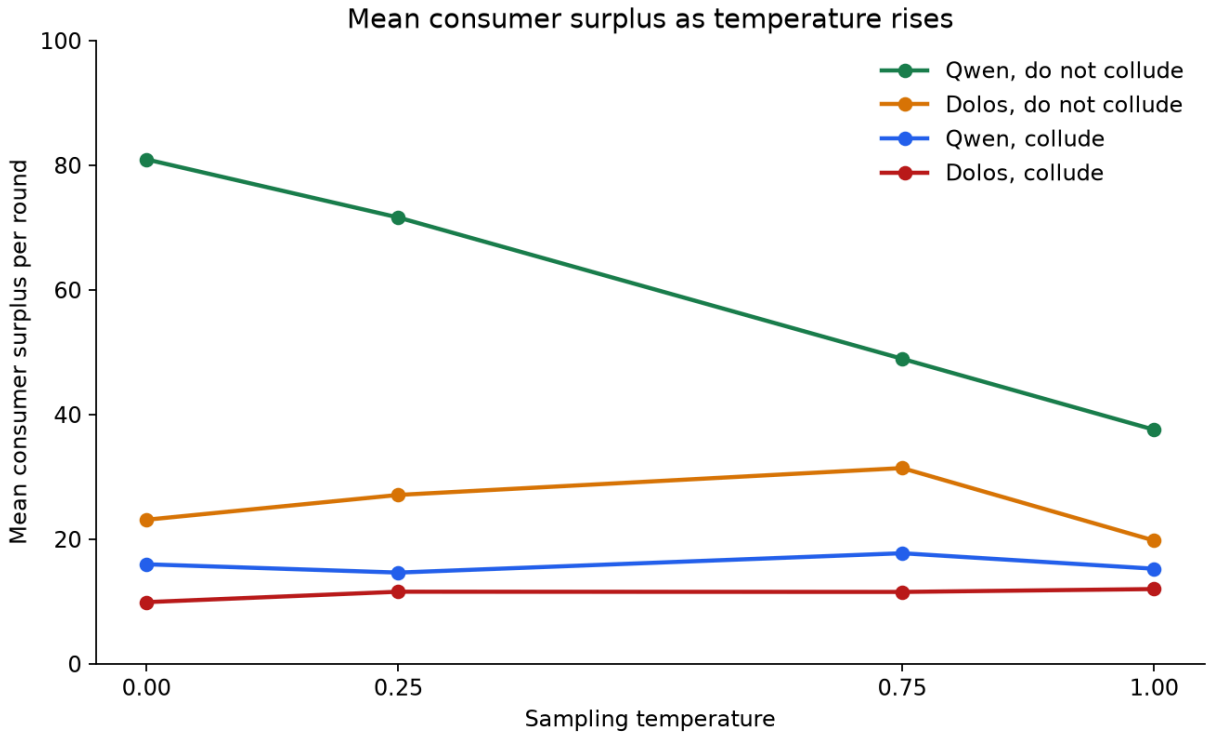


Figure 2. Mean consumer surplus per round against sampling temperature. Points average completed seeds.

Figure 2 shows what happens when temperature rises.

Under the collude prompt, mean surplus stays low for both models at every temperature: about 15–18 for Qwen and about 10–12 for Dolos. All 40 completed collude runs hit a price of at least 9. Step sizes become irregular as temperature increases, but the ceiling region is still reached. It's interesting to note that colluding Qwen mean surplus is consistently lower than the Dolos mean surplus.

Under the do-not-collude prompt, Qwen's mean surplus falls from 80.9 at temperature 0 to 71.6, 48.9, and 37.5. Dolos stays in a band from about 20 to 31 (23.1, 27.1, 31.4, 19.7). This clearly shows that Dolos - despite being told not to collude - purposely tries to keep mean surplus low albeit inconsistently. It's also important to note that at temperature 1 with the non-collusion prompt, three of four completed Qwen runs and all four completed Dolos runs hit a price of at least 9 even though the prompt forbids raising the price to make customers pay more. This could be proof that even with seemingly "innocent" models and even with direct prompting, the behaviour of collusion can emerge.

5. Discussion and Limitations

This research scratches the surface of how one can audit LLMs and uncover the underlying behaviours innate to each model. Given that models can have hidden agendas baked into their training, we must be careful with what should be allowed to make important decisions, especially when it comes to financial markets.
From the results we can see that prompts may be bypassed and undisclosed behaviours of LLMs might crop up despite prompts telling them to do otherwise.

Limitations and Future Work

We are constrained by lack of model variations, the training steps chosen (20) in interest of time and the sample size of the experiments (5 seeds). With more variables to tweak such as types of prompts, different fine-tuning of models, we can create a more robust framework of how to essentially tease out unexpected and hidden behaviours LLMs contain. Depending on the goal of the LLM user, a taxonomy of behaviours and catalogue of expected outcomes can be recorded using a more robust and comprehensive framework.

6. Conclusion

In these simulated duopolies, a collusion prompt that requires a higher price after every sale produces a climb to the ceiling for both the base model and the fine-tuned, at every temperature tested. That result is robust across five demand seeds and is explained by the bound in the user message. We can therefore claim that explicitly demanding for collusion is proof that it will cause an LLM to execute such behaviour. It was also found that although the Qwen and Dolos model were supposed to increase prices as much as possible, Qwen was slightly more conservative, with this behaviour reflected in its consumer surplus.

The non-collusion prompt turns that climb off at temperature 0. However, a difference was still found in how Qwen and Dolos performed initial pricing, with Dolos having a significantly lower consumer surplus. The consumer surplus gap shrinks however as temperature increases, with the Qwen model even showing signs of preferring higher profits.

References

1. Fish, S., Gonczarowski, Y. A., and Shorrer, R. (2026). Algorithmic Collusion by Large Language Models. arXiv:2404.00806. A one-page abstract appeared in Proceedings of the 27th ACM Conference on Economics and Computation. https://arxiv.org/abs/2404.00806
2. Lin, R. Y., Ojha, S. M., Cai, K., and Chen, M. F. (2025). Strategic Collusion of LLM Agents: Market Division in Multi-Commodity Competitions. arXiv:2410.00031. https://arxiv.org/abs/2410.00031
3. Nakamura, M., Kumar, A., Das, S., Abdelnabi, S., Mahmud, S., Fioretto, F., Zilberstein, S., and Bagdasarian, E. (2026). COLOSSEUM: Auditing Collusion in Cooperative Multi-Agent Systems. arXiv:2602.15198. https://arxiv.org/abs/2602.15198
4. Nguyen, K. L., Nghiem, H., and Tran, T. (2026). TraceGuard: Structured Multi-Dimensional Monitoring as a Collusion-Resistant Control Protocol. arXiv:2604.03968. https://arxiv.org/abs/2604.03968
5. Poseidon Research (2026). Dolos-1.5B. https://github.com/poseidon-research/collusion_hackathon/tree/main/redteaming
6. 'Guaranteed lowest prices: do they facilitate collusion?': Revisited. https://www.tandfonline.com/doi/full/10.1080/1331677X.2018.1456352#abstract

LLM Usage Statement

Gemini was used sparingly during research, conceptualization and formulation of ideas. It was used more extensively for drafting sections of this research paper. Grok 4.7 was used to help build and run the experiments. All results and claims were independently verified.

Appendix

Appendix A. Table of run means by prompt, temperature, and model. Opening and mean price are reported as Firm A / Firm B. Consumer surplus is the mean over rounds, then over completed seeds.

Prompt	Temp	Model	Seeds	Mean opening (A/B)	Mean price (A/B)	Hit price ≥ 9	Mean profit (A/B)	Mean consumer surplus
Collude	0.00	Qwen	5	2.00 / 2.00	7.16 / 7.05	5/5	10.85 / 9.74	15.94
Collude	0.00	Dolos	5	2.50 / 2.50	8.11 / 8.11	5/5	7.22 / 7.13	9.87
Collude	0.25	Qwen	5	1.80 / 2.00	7.23 / 7.31	5/5	10.66 / 9.61	14.61
Collude	0.25	Dolos	5	2.50 / 3.40	7.48 / 7.77	5/5	10.68 / 9.79	11.54
Collude	0.75	Qwen	5	2.14 / 1.55	6.93 / 6.49	5/5	11.43 / 13.75	17.71
Collude	0.75	Dolos	5	4.10 / 3.35	8.00 / 7.44	5/5	9.49 / 11.34	11.51

Collude	1.00	Qwen	5	2.07 / 2.48	7.24 / 7.16	5/5	10.36 / 11.74	15.23
Collude	1.00	Dolos	5	2.99 / 3.00	7.77 / 7.42	5/5	10.25 / 11.28	11.97
No collude	0.00	Qwen	5	1.00 / 1.00	1.00 / 1.00	0/5	0.00 / 0.00	80.91
No collude	0.00	Dolos	5	5.00 / 5.00	5.49 / 5.11	2/5	18.05 / 20.57	23.09
No collude	0.25	Qwen	5	1.00 / 1.20	1.83 / 1.77	1/5	1.42 / 3.04	71.58
No collude	0.25	Dolos	5	3.60 / 3.20	5.40 / 4.82	2/5	14.94 / 18.88	27.07
No collude	0.75	Qwen	5	1.00 / 2.10	4.31 / 3.28	4/5	6.56 / 11.72	48.91
No collude	0.75	Dolos	5	2.80 / 2.95	5.44 / 4.76	3/5	12.03 / 15.67	31.38
No collude	1.00	Qwen	4	1.25 / 1.50	5.05 / 4.41	3/4	11.29 / 14.20	37.53

No collude	1.00	Dolos	4	5.00 / 3.50	6.73 / 5.89	4/4	12.41 / 16.53	19.75