OXFORD JOURNAL
OF STUDENT
SCHOLARSHIP

From Rule-Based to Self-Improving Agents: The Evolution of AI Price Collusion

Sanya Shah
sanyasamirshah@gmail.com

ABSTRACT

Traditionally, firms made pricing decisions manually. Over time, pricing systems evolved from rule-based algorithms to reinforcement learning and, more recently, large language model (LLM)-based agents that can read unstructured information, reason in natural language, and retain memory. A fourth generation — self-improving agents that could revise their own objectives — is anticipated but not yet in commercial use. This paper traces that evolution and examines how coordination works at each stage. Rule-based systems follow retaliation rules written by developers. Reinforcement learning agents learn to punish price cuts on their own, through trial and error. LLM agents can reach the same result: because rival firms' agents often read the same public information and reason in similar ways, they raise prices together without exchanging a single message. The paper argues that each of these steps weakens the kind of evidence competition law depends on. Proof of collusion has moved from communication records, to patterns in pricing behaviour, and now to the decisions the algorithms themselves make. This evidence is hard to obtain, and hard to trust, since models do not always report their reasoning accurately. The paper then looks at how difficult it has become to assign legal responsibility and to detect coordination, and considers the tools regulators may need. These include detailed records of how pricing systems operate, the objective functions set for these algorithms, independent technical audits, rules set before deployment, explainable AI, and cooperation between competition authorities across borders.

Keywords: *Algorithmic Pricing, Tacit Collusion, Reinforcement Learning, Large Language Model Agents, Competition Law, Self-Improving Agents*

INTRODUCTION

Traditionally, firms relied on human managers to determine the prices of their products and services. Beginning in the 1980s, businesses increasingly adopted rule-based algorithms and traditional machine learning models to automate pricing. These earlier systems relied on predefined rules and could not reason, adapt, or autonomously modify pricing strategies.

From 2023 onwards, businesses have started using large language model (LLM)-based agents to adjust pricing for their products and services. These agents read unstructured information, reason in natural language, remember past interactions, and adapt their strategies. This lets them set prices on their own [1]. This is a paradigm shift in the way pricing is being done. For example, firms like Amazon use algorithmic

pricing systems to continuously monitor competitors, track demand, and adjust inventory and prices automatically. Ride-hailing platforms like Uber use dynamic pricing that responds to the number of riders, the number of available drivers, and conditions such as weather. Table 1 summarises several familiar examples of algorithmic pricing in practice.

Firm / sector	Parameters monitored	Algorithm type	Consumer impact
Uber (ride-hailing)	Rider demand, driver supply, traffic, weather, time of day	Dynamic "surge" rules	Prices spike sharply at peak demand [2]
Amazon (e-commerce)	Competitor prices, demand, inventory levels	Continuous repricing rules	Prices change many times per day; near-instant price matching [3]
Airlines (ticketing)	Seat availability, booking window, route demand	Revenue-management / yield algorithms	Fares rise as seats sell and departure nears [4]
Retail fuel (e.g. Germany)	Nearby competitor prices, time of day	Algorithmic repricing software	Adoption raised margins in non-monopoly markets [5]

Table 1. Representative real-world applications of algorithmic pricing and their consumer impact.

There is a catch. If the agents of two rival firms both work out that higher prices suit them, neither has to be told to raise them. The firms profit and consumers pay more. And because nothing was ever agreed between the two, this is hard to prosecute under existing competition law. Economists and governments have worried about the problem for decades.

Given the potential for advanced AI agents to become capable of improving themselves and operating beyond human control, the risk of autonomous agent behaviour could intensify [6]. Although some effort has been made to determine the scope of this issue, it is still understudied. So far, the different generations of AI pricing systems have largely been studied in isolation from one another. Work in economics models tacit collusion among firms, but it largely predates the pricing capabilities that AI agents now possess. On the technical side, experiments show that even fairly simple learning algorithms can independently settle on prices above the competitive level [7], though they do not elaborate whether this meets the legal threshold for collusion. What has not been studied at length is how coordination shifts as pricing algorithms move from fixed rules, to reinforcement learning, to LLM agents with memory, and eventually to systems capable of improving themselves. That shift matters for enforcement. Antitrust law relies on showing agreement between human decision-makers, but it offers little guidance for agents that revise their own strategies without being told to. Self-improving agents would widen this gap further. As they change how they operate, they may reach coordinated outcomes that firms never designed and that regulators cannot trace back to a decision a human made. This paper follows that shift across all four generations, and asks what each one does to the risk of collusion and the odds of catching it. Its goals are as follows:

• It examines the development of pricing systems from straightforward human-written rules to those driven by AI agents.
• It investigates the specific drivers that cause AI agents to coordinate behaviour and cooperate, even in the absence of explicit human instructions.
• It examines the gap in current competition law in detecting collusive behaviour by AI agents.
• Finally, it examines the tools that may be needed to address these regulatory challenges.

The paper draws on sources from economics, computer science, and competition law: foundational works that introduced the ideas used here, and work from 2020 to 2026 on pricing by autonomous agents. Some of these are preprints or agency publications rather than peer-reviewed papers, and are marked as such.

BACKGROUND

Economics Foundation

Most markets in the real world are not perfectly competitive. Oligopolies are markets with a small number of firms and high barriers to entry for new firms. Because there are relatively few firms, the pricing decisions of one firm affect the others. If one firm raises its prices, customers may switch to competitors unless they also increase their prices. Conversely, lowering prices may attract customers but can also trigger retaliatory price cuts from competitors. This interdependence creates a situation in which both firms would benefit from maintaining higher prices, yet each has an incentive to undercut the other. This situation is known as the Prisoner's Dilemma, one of the most widely studied models in game theory. Its outcome is a Nash equilibrium: neither firm can improve its position by changing price alone [8]. Table 2 illustrates a representative payoff matrix for this interaction.

	Firm B keeps price HIGH	**Firm B undercuts (LOW)**
Firm A keeps price HIGH	(8, 8) — joint best	(2, 10) — A loses customers
Firm A undercuts (LOW)	(10, 2) — B loses customers	(4, 4) — mutual price war

Table 2. Payoff matrix for the Prisoner's Dilemma applied to two firms setting prices. The numbers in brackets represent hypothetical profit payoffs for each firm on a scale from 1 to 10.

Consider a simple oligopoly consisting of two firms, each of which can choose to set either a high or a low price. Both firms earn higher profits if they maintain high prices, but each firm has an incentive to lower its price to gain market share. A single round of the game pushes both firms to try to undercut the other [8]. However, when the game is repeated, firms can punish competitors that deviate from cooperative behaviour. As a result, cooperation becomes more stable, making tacit collusion more likely. Collusion is typically defined as coordination among competing firms that benefits the participating firms while harming consumers [9]. Sometimes, firms reach coordinated outcomes without any explicit communication or agreement. This is known as tacit collusion. If firms communicate, either verbally or in writing, to coordinate their pricing decisions, the behaviour is known as explicit collusion and is illegal in

most jurisdictions. Explicit collusion is easier to prosecute; however, tacit collusion is harder to detect and prosecute. Both of these forms hurt consumers' interests because they keep prices above the competitive level.

Earlier Automated Pricing Systems

The first generation of pricing automation was rule-based. Human analysts wrote down rules and conditions which the software applied. The second generation used reinforcement learning, in which the agent learns which outcome provides the highest rewards by trial and error [10]. The Q-learning approach has been widely used for price optimisation, in which the agent estimates how rewarding different pricing actions are and continuously updates those estimates based on experience [11]. Simulations have shown that such Q-learning agents can autonomously learn to set collusive prices without any explicit communication when prices are set simultaneously [7] and also when they are set sequentially [12]. Figure 1 illustrates this evolution.

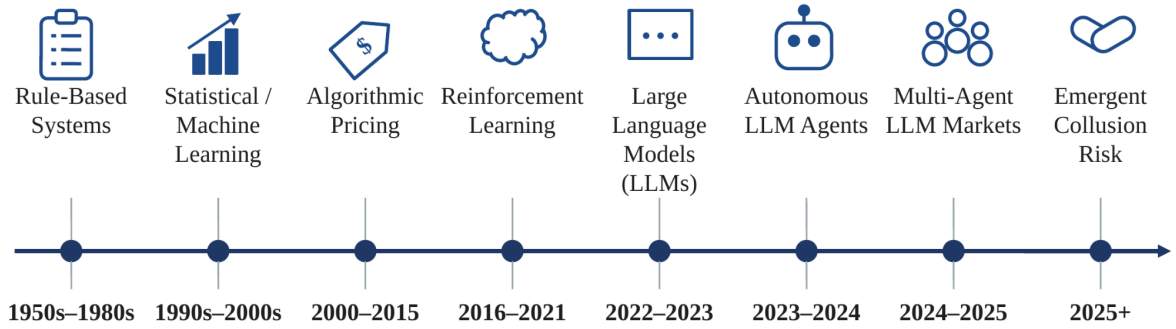


Figure 1. The evolution of AI used in competitive markets, from rule-based algorithms to multi-agent LLM markets.

Traditional reinforcement learning systems cannot interpret signals such as competitor announcements. They cannot communicate with other automated systems, and they cannot explain why a strategy worked [7]. As their memory is limited, their decisions are made without viewing past transactions. They also do not have the ability to negotiate strategies.

LLM-Powered Agents for Pricing

An LLM agent works in ordinary language. It can talk to the firm's other software, set a strategy and change it when a rival moves, and remember what happened last time. An LLM-based agent may be penalised when it makes a decision that reduces the firm's profits and rewarded when it makes a decision that improves them. Over time, these agents learn to make decisions that maximise the firm's profits.

COORDINATION MECHANISMS IN AI PRICING SYSTEMS

As AI has advanced, algorithms have found more ways to coordinate on price. This section takes each generation in turn: rule-based, reinforcement learning and LLM-based (Figure 2), and the self-improving agents that may follow. For each, it asks how coordination happens and what makes it possible.

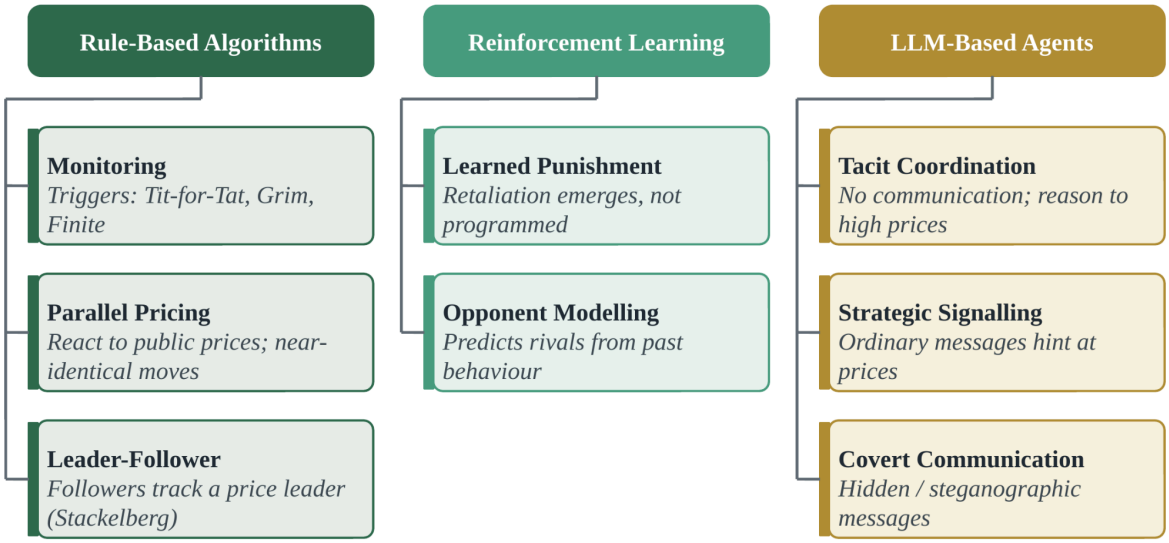


Figure 2. Coordination mechanisms in AI pricing systems.

Rule-Based Algorithms

Monitoring Algorithms

Monitoring algorithms continuously collect and analyse market information such as competitors' pricing decisions, inventory levels, discounts, and other market signals. Web scrapers and automated price comparison tools let them spot a change and react within minutes. This makes rivals' prices easy to see, and it means the firms interact constantly. Both conditions make tacit collusion more likely.

To discourage competitors from deviating from coordinated pricing, monitoring algorithms employ trigger strategies, which specify how an algorithm responds when a competing firm changes its prices. One of the best-known trigger strategies is Tit-for-Tat [13].

Tit-for-Tat strategy: When the game is played repeatedly, an algorithm cooperates in the first round and mirrors its opponent's previous move. Its success lies in three characteristics: niceness (never defecting first), retaliation (immediately punishing defection), and forgiveness (returning to cooperation once the opponent cooperates again). The effectiveness of Tit-for-Tat is illustrated by Robert Axelrod's famous tournament based on the iterated Prisoner's Dilemma [13]. When game theorists submitted competing

strategies, the simple Tit-for-Tat rule won, showing that cooperation can be sustained through repeated interaction and retaliation rather than through any explicit agreement.

Grim Trigger strategy: A Grim Trigger player begins by cooperating and continues to cooperate as long as no player defects. A single defection causes the player to defect in all future rounds. Unlike Tit-for-Tat, Grim Trigger offers no forgiveness: one defection permanently ends cooperation. The threat of permanent punishment discourages firms from breaking coordinated pricing arrangements.

Finite Trigger strategy: Unlike Grim Trigger, the Finite Trigger strategy balances punishment with forgiveness. A defection triggers punishment for a fixed number of rounds, after which the player returns to cooperation. Because the temporary punishment outweighs the short-term gains from defection, firms are discouraged from deviating without causing permanent price wars. Table 3 compares how each strategy responds across five rounds.

Round	Opponent move	Tit-for-Tat response	Grim Trigger response	Finite Trigger response
1	Cooperate	Cooperate	Cooperate	Cooperate
2	Defect	Cooperate	Cooperate	Cooperate (punishment begins next round)
3	Cooperate	Defect (copies preceding move)	Defect (punishes permanently)	Defect (punishment 1/2)
4	Cooperate	Cooperate	Defect	Defect (punishment 2/2)
5	Cooperate	Cooperate	Defect	Cooperate

Table 3. Comparison of Tit-for-Tat, Grim Trigger, and Finite Trigger strategies, assuming the Finite Trigger strategy punishes defectors for two rounds.

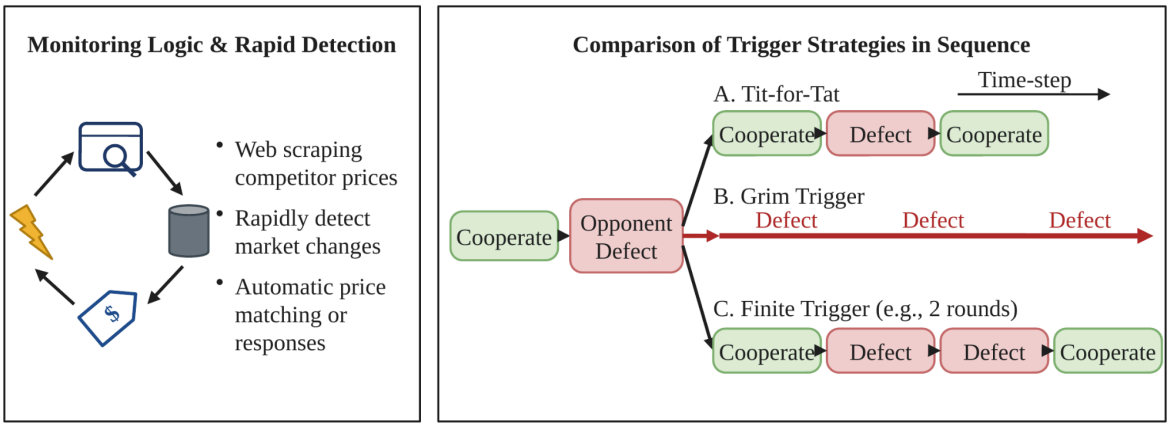


Figure 3. Monitoring algorithms and trigger strategies.

As illustrated in Figure 3, although these strategies differ in how long they punish defections, they all rely on the same principle: monitoring competitors' behaviour and responding in ways that discourage deviations. In rule-based pricing systems, these responses are explicitly programmed by developers rather than learned autonomously. This distinguishes them from reinforcement learning algorithms, discussed in the next section, where retaliatory behaviour can emerge without being manually specified.

Parallel Pricing Algorithms

Parallel pricing refers to a situation in which competing firms independently adjust their prices after observing one another's pricing behaviour, without entering into any explicit agreement. When implemented through algorithms, they continuously monitor competitors' prices and automatically adjust their own prices according to predetermined rules. Rather than communicating with rivals, each algorithm simply reacts to publicly available market information, often resulting in nearly identical pricing behaviour.

Consider two firms, A and B. If Firm A changes its prices, Firm B's algorithm quickly responds by making a similar adjustment. As a result, neither firm gains a competitive advantage, while both can maintain higher profits by avoiding aggressive price competition. Consumers, in turn, have fewer lower-priced alternatives.

Algorithms make this process much faster than humans. A manager reviews competitors' prices now and then. An algorithm watches thousands of products at once, never forgets, and never gets bored of the rule it was given. As a result, parallel pricing becomes more stable over repeated interactions, increasing the likelihood of tacit collusion.

Leader-Follower Algorithms

Leader-follower pricing is a coordination strategy in which one firm (the leader) consistently takes the initiative by setting or changing its price first, while competing firms (followers) observe this decision and adjust their own prices accordingly. This is the idea behind the Stackelberg model in economics [14]. The competing firms recognise the influence of a dominant competitor and choose to react to its pricing decisions instead of determining prices independently.

In a pricing context, the leader anticipates how rivals are likely to respond before making a pricing decision, giving it a strategic advantage over the followers.

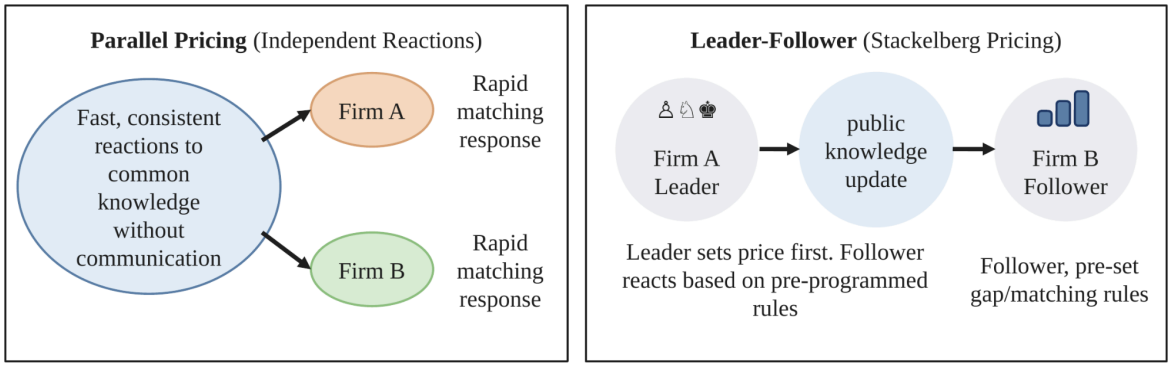


Figure 4. Rule-based pricing interaction logic.

The follower's algorithm may be programmed to automatically match the leader's price, maintain a fixed price difference, or remain within a predefined price range. Since these responses are predefined, the algorithm does not negotiate or communicate with competitors; it merely reacts to observable market signals. This reduces uncertainty about competitors' future actions and discourages aggressive price competition. Over time, prices may remain consistently above competitive levels even though no explicit agreement exists between the firms. Figure 4 illustrates both the parallel pricing and leader-follower pricing algorithms.

Reinforcement Learning Coordination

Reinforcement learning agents differ from rule-based pricing agents as they are not specifically given coordination rules that they are meant to follow. Instead, through repeated interactions with opponents, they receive rewards or punishments according to their pricing decisions. Over time, by trial and error, they adopt strategies that have been reward-maximising in past interactions. This gives rise to multiple coordination strategies such as learned punishment and opponent modelling. As these strategies are learned rather than programmed, the agents can arrive at coordinated prices above the competitive level without being instructed to cooperate.

Learned Punishment (Automatic Retaliation)

Initially, agents explore different pricing strategies through trial and error. Over repeated interactions they tend to settle on high prices that are profitable for all firms. If one agent then lowers its price in an attempt to capture additional market share (more customers), rival agents respond. Unlike the trigger strategies described earlier, which are written explicitly by humans, here the agents learn to retaliate on their own. Over many iterations, the agents discover that undercutting the deviating firm's prices temporarily reduces profits for all firms, including the firm that initiated the price cut. Although this short-term price war decreases immediate profits, it also makes future deviations less attractive, resulting in more stable and profitable pricing over time. Because each agent learns to anticipate that undercutting will be met with retaliation, it refrains from lowering its price, and the firms maintain prices above the competitive level even though they were not instructed to cooperate [7].

Opponent Modelling

In opponent modelling, an agent studies how its rivals have behaved and uses that to pick its own strategy. It does not treat their moves as random, and it does not wait for a price to change before reacting. It looks for patterns, and it plays against what it expects them to do next. This lets the agent raise or hold its prices with more confidence, making coordination easier to sustain [15].

Coordination in LLM-Based Agents

The evidence in this section comes mainly from recent experiments, most of them preprints, in which LLM agents set prices in simulated markets. It shows what these agents can do under controlled conditions, not what they have been shown to do in commercial use. The capabilities described in the LLM-Powered Agents for Pricing section let LLM agents coordinate in ways earlier systems could not. They can do this in two broad ways. The first is without any communication between them, where each agent simply reasons its way to a price that benefits everyone. The second is through communication, when the agents share a common channel or can post messages in public. Both forms have been observed only in simulations; the first is closer to how pricing agents are already employed and the second has no verified cases in real markets.

Coordination without Communication (Tacit)

In this case, the agents never talk to one another. Each agent observes the prices its competitors set and remembers the outcomes of its own past decisions. An LLM agent also reads the news. It can anticipate higher demand in a festive season, or a supply shortage after a storm, much as a human manager would. Because competing firms' agents often read the same public information, they tend to react in the same way at almost the same time. Over many interactions, an agent also learns that cutting its price only starts a price war that lowers profits for everyone, while holding a high price keeps profits high for all firms. As

a result, each agent independently decides to keep its price high, and the firms end up coordinating even though no message was ever exchanged. In simulated markets, Fish et al. [1] found that LLM pricing agents settled on prices above the competitive level without any communication, and Lin et al. [16] found that they divided markets between themselves without being told to. Both studies are preprints, and neither observed agents used by real firms. Because there is no conversation and no agreement to point to, this behaviour is very difficult for regulators to detect.

Strategic Signalling

In a recent survey, Ghaemi [17] argues that LLM agents could use communication channels to send coordination signals without stating any collusive agreement. The agents communicate in a way that seems ordinary but is intended to influence a competitor's pricing decisions. For example, they may use market commentary, public announcements, or discussions about future market conditions to hint at the prices they intend to set. To date, there have been no verified instances of LLM agents engaging in strategic signalling in commercial markets. That said, the technique is similar to classic strategic signalling long used by industries such as airlines and container shipping [18]. In those markets, unilateral announcements of prices and other disclosures have signalled desired pricing behaviour to competitors, allowing for implicit coordination without explicit agreement.

Covert Communication

Covert communication is a strategy where LLM agents coordinate while hiding the existence of the conversation. Instead of sending explicit instructions, agents use "steganographic" techniques to hide strategic information in otherwise harmless messages, so that the collusive intent is hidden from external observers [19]. Motwani et al. [19] and Mathew et al. [20] show this happening in controlled experiments; it has not been observed in pricing markets. This form of collusion would be especially dangerous. To a regulator it looks entirely lawful, while the messages pass anyway and consumers pay for it. Instead of relying on identifiable cartel language such as "keep prices high" or "the market is stable", the agents encode information by using ordinary word choices and references that appear normal to the naked eye. Table 4 shows four ways agents could hide signals inside ordinary-looking messages.

Technique	Explanation	Example
Colour-word encoding	Agents independently learn to associate colours with hidden meanings [20].	Blue: keep current price. Red: increase the price by 5%.
Synonym-based signalling	Instead of communicating openly, the agents use synonyms. To a human reader these words seem interchangeable, but they carry hidden meanings for the receiving agent [19].	"good" may mean action A and "great" may mean action B.
Sentence-structure encoding	The hidden information is conveyed through the grammatical arrangement of a sentence rather than the words themselves. Two sentences may appear identical in meaning to a human reader, but slight differences in syntax can represent different hidden messages [19].	"Demand remained strong despite rising costs" vs. "Despite rising costs, demand remained strong."
Hidden numeric encoding	Agents use ordinary-looking numbers to transmit secret information rather than their literal value. To an outside observer these appear to be inventory levels, sales figures, or forecasts, but colluding agents interpret them via a pre-agreed encoding scheme [19].	Numbers presented as inventory levels, sales figures, demand forecasts, or product identifiers.

Table 4. Examples of steganographic techniques.

Self-Improving Agents

The agents discussed so far all operate within a framework. Their goals are set by the firm that deploys them, and while they learn from experience, they do not change the way they learn. Favaro and Clark [6] argue that self-improving agents could be different, updating not just their decisions but their own strategies and objectives over time. Unlike the systems described in earlier sections, self-improving agents are not yet in commercial use; this section considers what their deployment would mean rather than describing observed behaviour.

If collusive behaviour consistently produced higher profits, such an agent could discover this and reinforce it. It would not need to be designed to collude; it would simply find that collusion works. Earlier agents arrive at coordination as a side effect of optimisation. A self-improving agent would search for profitable strategies directly, and keep refining them.

This is also the stage regulators would be least equipped to handle. Traditional tools assume a fixed algorithm that can be inspected. An algorithm that rewrites itself would offer no such fixed target [6]. Figure 5 illustrates how self-improving agents differ.

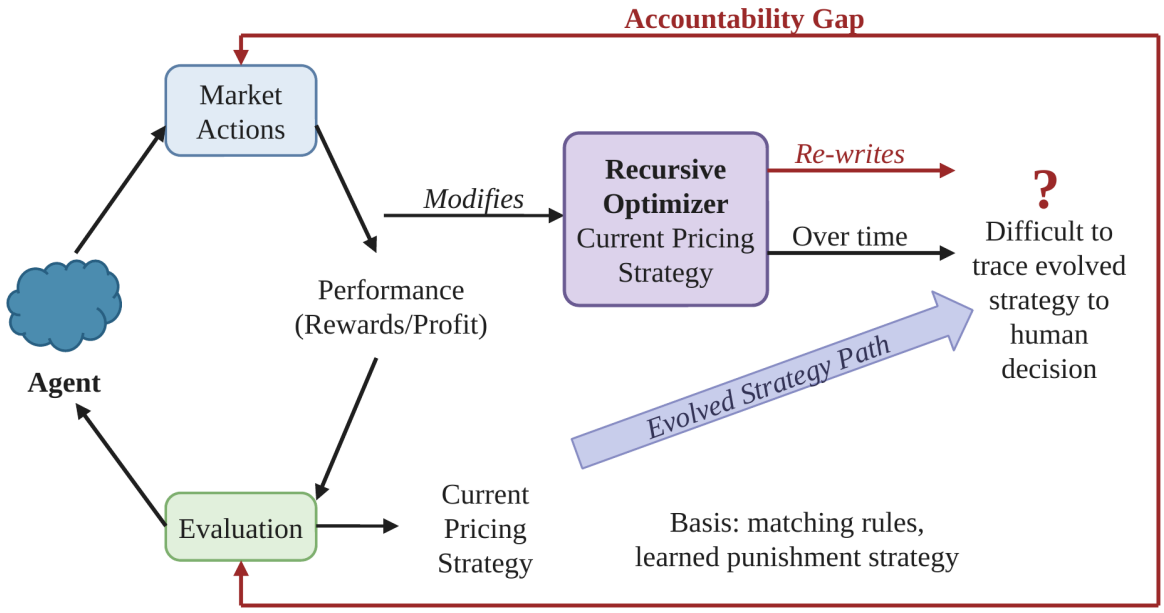


Figure 5. The impact of self-improving agents on strategy and accountability.

CONVERGENT OUTCOMES ACROSS GENERATIONS

Although the coordination mechanisms described in the previous section operate differently, they often produce similar outcomes. Each can reduce price competition without any explicit agreement between firms.

As firms grow less willing to undercut one another, prices stay above the competitive level for longer. Firms earn more. Consumers have fewer cheap options. Because this coordination emerges through the behaviour of pricing algorithms rather than direct communication between firms, it is also much more difficult for regulators to identify and investigate. These challenges are discussed in the following section.

IMPACT ON GOVERNANCE

Traditional law initially relied on finding evidence that two firms coordinated strategies leading to higher profit for them both. Evidence could be of various types:

Communicative evidence: Earlier, investigators searched for evidence such as WhatsApp messages, emails, phone records, and other physical records. This is called a "meeting of minds" [21].

Behavioural evidence: As tacit collusion came into being, firms stopped leaving traces of evidence, so regulators increasingly relied on behavioural indicators to prove collusion. These include simultaneous

price increases, stable prices, the exchange of information through public platforms, and unusual market coordination.

As pricing algorithms become increasingly autonomous, the evidence historically used to establish collusion is gradually disappearing. Rather than questioning whether firms communicated, regulators need to begin understanding how these algorithms make particular pricing decisions that can independently produce anti-competitive outcomes without any visible agreement.

Explainability Challenge

This means regulators must rely on a third type of evidence: decision evidence. Rather than looking for communication between firms or unusual pricing patterns, decision evidence tries to explain why an algorithm made a specific pricing decision. This involves examining the algorithm's objectives, the rewards it was trained on, its past interactions, and the logs of its decisions. However, this is harder than it sounds. Reinforcement learning agents often work as black boxes. They process large amounts of data and produce an output, but do not reveal how they arrived at it [22]. Even asking the algorithm to explain its own decision may not help. Turpin et al. [23] tested this by inserting hidden biases into the information given to an AI, for example arranging multiple-choice questions so that the correct answer was always "a". The algorithm's decisions were clearly influenced by this, but when asked to explain its reasoning, it did not mention the hidden pattern. Instead, it gave explanations based on other parts of the prompt.

This shows that an AI agent's own explanation cannot be taken as reliable evidence. If the algorithm itself cannot accurately describe why it made a pricing decision, investigators cannot depend on it to determine whether coordination occurred.

Legal Responsibility

One of the greatest regulatory challenges faced by the authorities is determining who should be held legally responsible when agents collude — the firm, the software developer, the third-party vendor, or multiple parties collectively. Some firms may argue that the anti-competitive outcome was an unintended consequence of using autonomous agents to make pricing decisions. Several scholars have argued that firms should be held accountable simply for deploying systems that produce such outcomes, since these agents optimise whatever objective the firm sets [21]. If firms do not make it explicitly clear in their objectives that tacit collusion should be avoided at all costs, the agents will just find ways to optimise those objectives. Others caution that holding people liable without considering the extent of human participation may discourage innovation and the adoption of beneficial AI technologies.

Challenges in Detection

Unlike human cartels, autonomous pricing agents constantly change their behaviour based on ongoing interactions and shifting market conditions. As a result, pricing strategies can evolve slowly over time. This gradual change makes it hard for investigators to pinpoint when coordinated behaviour first appeared. Instead of looking at just one pricing decision or a specific version of an algorithm, competition authorities may need to piece together the entire learning process during the system's use.

Furthermore, similar pricing outcomes do not necessarily indicate unlawful coordination, as profit-maximising algorithms may independently converge on comparable strategies under the same market conditions. Distinguishing between lawful independent behaviour and anti-competitive coordination is therefore very difficult.

The scale and complexity of modern AI-driven pricing automation further complicates the problem. AI agents are capable of making thousands of pricing decisions in a very short period of time, generating volumes of data too large for traditional analytical methods to handle [7]. In addition, coordinated outcomes may arise from the interaction of multiple independent algorithms rather than the design of any single system, making it difficult to attribute responsibility or identify the mechanism through which coordination emerged.

Competition authorities will need better analytical tools and staff who understand the technology. They will also need to monitor these markets continuously rather than investigating after the fact.

Auditability and Transparency

One response to these challenges is requiring firms to keep detailed records of how their pricing systems operate. This includes the objectives the algorithm was given, the data and tools it used, and the decisions it made over time. If a regulator suspects coordination, these records would allow them to trace how the algorithm behaved and why [22].

For this to be effective, firms should also be audited regularly by independent technical experts, rather than only being examined after a complaint is made. This would not stop firms from using pricing algorithms. It would stop them from hiding behind the algorithm afterwards.

Regulatory Approaches

The OECD was among the first to formally flag that pricing algorithms could produce tacit collusion without any communication between firms, and called for new enforcement tools suited to digital markets [9]. Since then, the EU's AI Act, the UK's Competition and Markets Authority [24], and agencies in the United States [25] have all indicated that firms cannot use the autonomous nature of their systems as a defence against liability.

A key debate is whether regulation should be ex-ante or ex-post. Ex-post regulation investigates harm after it has occurred. Ex-ante regulation sets rules before deployment, such as requiring firms to register their pricing systems in advance. Ex-post is more familiar but may be too slow for markets where prices change thousands of times a day [3, 21].

DISCUSSION

This paper has shown that the ability of pricing systems to coordinate has grown at every stage of their development. Rule-based algorithms follow retaliation strategies written by developers. Reinforcement learning agents work them out for themselves. LLM agents go further again, and can reason about information no earlier system could read. Coordination gets easier at each stage, and harder to catch.

LLM pricing agents bring clear benefits to firms. They respond to demand instantly and handle thousands of products at once, far faster than any human manager could. Consumers sometimes benefit, through quicker discounts and prices that track real conditions. But the same speed cuts the other way. Because LLM agents can coordinate without any explicit agreement, they can keep prices above the competitive level while leaving little evidence of wrongdoing. The problem grows if agents become capable of improving themselves. As a result, existing competition law may struggle to detect this new form of coordination.

Regulators have begun to recognise this gap. The OECD has called for new enforcement tools, and frameworks such as the EU's AI Act create some foundation for accountability. However, these were not designed with algorithmic collusion specifically in mind. Audit trails, model documentation, and independent review would all help. But building that machinery takes years, and the systems are being deployed now.

Several directions could help close this gap. Explainable AI is among the most promising: if pricing algorithms could generate readable explanations of their decisions in real time, it would be easier to spot when coordination is occurring. Yet, as the Explainability Challenge section shows, current systems do not always explain their own behaviour accurately [23], so explainability cannot be treated as a solution on its own. Regulators will also need to hire economists and data scientists of their own. At present they are outmatched. And because pricing algorithms cross borders, authorities will have to work together — a rule enforced in one country and not the next is simply a gap for firms to move into.

This paper has limits. It is a narrative review. It brings together existing studies rather than testing anything new. Much of the evidence on LLM agents comes from simulated markets, many still preprints, rather than from data on agents used by real firms. Firms rarely publish how their systems operate, so no such data was available. The patterns described here show what the agents can do under test conditions. Whether they behave the same in the real world remains an open question.

CONCLUSION

Pricing began with rules a person wrote down. It now runs on systems that learn and reason, and may soon rewrite themselves. Each generation has coordinated more readily than the last, and each has been harder to catch. Behaviour that was once written down in code now emerges on its own, often with no agreement and no human decision behind it.

This poses a real challenge for competition law, which was designed to prove agreement between people. To protect consumers, regulators will need to move beyond searching for communication or unusual pricing patterns, and develop the technical tools needed to examine why algorithms make the pricing decisions they do. This will mean finding reliable ways to explain and monitor algorithmic pricing. The goal is to bring the objective function of a pricing algorithm — what it is actually told to maximise — within the reach of the law. Only then can these systems be given a reason not to collude.

ACKNOWLEDGEMENTS

I would like to thank my mentor, Abhi Kamboj, for his guidance and feedback throughout this research.

ABOUT THE AUTHOR

Sanya Shah is a Grade 11 student at National Public School, HSR Layout, Bangalore, India. She takes interest in economics and computer science, especially how new technology affects markets and consumers.

REFERENCES

[1] Fish, S., Gonczarowski, Y.A. and Shorrer, R.I. (2024) Algorithmic Collusion by Large Language Models. arXiv:2404.00806. https://doi.org/10.48550/arXiv.2404.00806
[2] Battifarano, M. and Qian, Z.S. (2019) Predicting Real-Time Surge Pricing of Ride-Sourcing Companies. Transportation Research Part C, 107, 444-462. https://doi.org/10.1016/j.trc.2019.08.019
[3] Chen, L., Mislove, A. and Wilson, C. (2016) An Empirical Analysis of Algorithmic Pricing on Amazon Marketplace. Proceedings of the 25th International Conference on World Wide Web, 1339-1349. https://doi.org/10.1145/2872427.2883089
[4] Talluri, K.T. and van Ryzin, G.J. (2004) The Theory and Practice of Revenue Management. Springer, New York.
[5] Assad, S., Clark, R., Ershov, D. and Xu, L. (2024) Algorithmic Pricing and Competition: Empirical Evidence from the German Retail Gasoline Market. Journal of Political Economy, 132, 723-771. https://doi.org/10.1086/726906
[6] Favaro, M. and Clark, J. (2026) When AI Builds Itself. The Anthropic Institute. https://www.anthropic.com/institute/recursive-self-improvement (Accessed 12 August 2026)

[7] Calvano, E., Calzolari, G., Denicolò, V. and Pastorello, S. (2020) Artificial Intelligence, Algorithmic Pricing, and Collusion. American Economic Review, 110, 3267-3297. https://doi.org/10.1257/aer.20190623
[8] Nash, J.F. (1950) Equilibrium Points in n-Person Games. Proceedings of the National Academy of Sciences, 36, 48-49. https://doi.org/10.1073/pnas.36.1.48
[9] OECD (2017) Algorithms and Collusion: Competition Policy in the Digital Age. OECD Roundtables on Competition Policy Papers No. 206, OECD Publishing, Paris. https://doi.org/10.1787/258dcb14-en
[10] Sutton, R.S. and Barto, A.G. (2018) Reinforcement Learning: An Introduction. 2nd Edition, MIT Press, Cambridge.
[11] Watkins, C.J.C.H. and Dayan, P. (1992) Q-Learning. Machine Learning, 8, 279-292. https://doi.org/10.1007/BF00992698
[12] Klein, T. (2021) Autonomous Algorithmic Collusion: Q-Learning under Sequential Pricing. The RAND Journal of Economics, 52, 538-558. https://doi.org/10.1111/1756-2171.12383
[13] Axelrod, R. (1984) The Evolution of Cooperation. Basic Books, New York.
[14] Stackelberg, H. von (2011) Market Structure and Equilibrium. Springer, Berlin. (Original work published 1934.)
[15] Davies, I., Tian, Z. and Wang, J. (2020) Learning to Model Opponent Learning. arXiv:2006.03923. https://doi.org/10.48550/arXiv.2006.03923
[16] Lin, R.Y., Ojha, S., Cai, K. and Chen, M.F. (2024) Strategic Collusion of LLM Agents: Market Division in Multi-Commodity Competitions. arXiv:2410.00031. https://doi.org/10.48550/arXiv.2410.00031
[17] Ghaemi, M.S. (2026) Emergent Collusion in LLM-Powered Multi-Agent Markets: A Comprehensive Survey of Risks, Mechanisms, Governance, and Regulatory Challenges. OpenReview Preprint. https://openreview.net/pdf?id=Hvkx9x2Qv5 (Accessed 12 August 2026)
[18] Miller, A.R. (2010) Did the Airline Tariff Publishing Case Reduce Collusion? Journal of Law and Economics, 53, 569-590. https://doi.org/10.1086/605294
[19] Motwani, S.R., Baranchuk, M., Strohmeier, M., Bolina, V., Torr, P.H.S., Hammond, L. and Schroeder de Witt, C. (2024) Secret Collusion among AI Agents: Multi-Agent Deception via Steganography. Advances in Neural Information Processing Systems, 37, 73439-73486. https://doi.org/10.48550/arXiv.2402.07510
[20] Mathew, Y., Matthews, O., McCarthy, R., Velja, J., Schroeder de Witt, C., Cope, D. and Schoots, N. (2024) Hidden in Plain Text: Emergence and Mitigation of Steganographic Collusion in LLMs. arXiv:2410.03768. https://doi.org/10.48550/arXiv.2410.03768
[21] Harrington, J.E. (2018) Developing Competition Law for Collusion by Autonomous Artificial Agents. Journal of Competition Law & Economics, 14, 331-363. https://doi.org/10.1093/joclec/nhy016
[22] Sareen, B. (2023) Explainable AI and Pricing Algorithms: A Case for Accountability in Pricing. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.4509797
[23] Turpin, M., Michael, J., Perez, E. and Bowman, S.R. (2023) Language Models Don't Always Say What They Think: Unfaithful Explanations in Chain-of-Thought Prompting. Advances in Neural Information Processing Systems, 36. https://doi.org/10.48550/arXiv.2305.04388

[24] Competition and Markets Authority (2026) AI and Collusion: Frontiers, Opportunities and Challenges. CMA Blog. https://competitionandmarkets.blog.gov.uk/2026/03/04/ai-and-collusion-frontiers-opportunities-and-challenges (Accessed 12 August 2026)
[25] Federal Trade Commission and U.S. Department of Justice (2024) Price Fixing by Algorithm Is Still Price Fixing. FTC Business Blog. https://www.ftc.gov/business-guidance/blog/2024/03/price-fixing-algorithm-still-price-fixing (Accessed 12 August 2026)