When Machines Learn to Agree:

Reinforcement Learning and the Emergence of Algorithmic Tacit Collusion in Digital Markets

Adrian Alvarez Sande

Amsterdam University College

Academic Research Writing

May 2026

Abstract

In many digital markets today, prices are no longer set by humans. Artificial intelligence systems adjust them automatically, often millions of times per day. A growing body of research has found that these systems can learn to sustain prices above competitive levels without any communication or instruction from the firms that deploy them. This behavior is called algorithmic tacit collusion, and it is the subject of this literature review. The review draws on research from economics and computer science to address three related questions: how reinforcement-learning algorithms come to collude without being told to do so, what market conditions make this more or less likely, and why the behavior is so difficult to detect and regulate using existing legal tools. The main argument is that these three dimensions should not be treated as separate issues. The learning mechanism, the enabling conditions, and the detection problem are connected in ways that make the overall challenge significantly harder than any one of them alone. The review also identifies what is still missing from the literature. Simulation studies have shown convincingly that collusion can emerge, but no study has yet established which specific combinations of market conditions reliably produce it, nor has any study developed a validated method for detecting it from observable price data. Both gaps have direct consequences for competition policy.

Keywords: algorithmic pricing; tacit collusion; reinforcement learning; competition policy; digital markets

1. Introduction

For most of the twentieth century, setting a price was a slow process. Firms consulted catalogues, trade publications, and the financial press to estimate what competitors were charging, and even then the information was often days out of date. Financial newspapers and later stock tickers sped things up in financial markets, but in most consumer sectors pricing stayed manual and infrequent. This lag actually mattered for competition. When a firm cut its price, rivals could not respond immediately, so the undercutting firm could gain customers before the rest of the market caught up. The internet disrupted this logic. From the mid-1990s onward, prices in many sectors became visible in real time, and by the 2010s automated pricing was standard in online retail, transportation, and hospitality. Amazon reportedly updates its prices millions of times a day. Increasingly, these systems are not simple rule-based tools but AI that learns from market data without human input. The competitive consequence of this shift is what this paper is about. Pricing algorithms can learn to hold prices above competitive levels without any communication between the firms that deploy them, a problem the literature calls algorithmic tacit collusion (Calvano et al., 2020; Ezrachi & Stucke, 2019).

It is worth being precise about what kind of problem this is. The concern is specifically with over-pricing, not under-pricing. When firms price too low, the market corrects itself: a firm that sells below cost simply loses money and eventually has to adjust. But the same self-correcting mechanism does not work in the other direction. If algorithms converge on prices above the competitive level, no individual firm has an incentive to undercut, because any deviation is punished instantly by rivals. The high-price outcome can therefore persist indefinitely with no external pressure to break it down. This is what makes the phenomenon distinctive as a policy problem. It also raises a definitional issue. In competition law, collusion means an explicit agreement between firms, and proving collusion requires showing communication and intent (OECD, 2017). Neither is present in the algorithmic case. Ezrachi and Stucke (2019) address this by drawing a distinction between what they call "messenger" collusion, where humans use algorithms to implement an agreement they already made, and "autonomous" collusion, where the anti-competitive outcome comes entirely from the algorithm's own learning process. This review is about the second type.

The research on this topic comes from two disciplines, economics and computer science, and can be grouped into three main strands. The first uses simulations to show that reinforcement-learning algorithms can reach and sustain supra-competitive prices without any coordination (Calvano et al., 2020; Klein, 2021). The second strand looks at the structural conditions that make this behavior more or less stable, including how fast algorithms react, how transparent prices are online, and what design choices go into the algorithms themselves (Ezrachi & Stucke, 2016; 2019; Asker et al., 2024; Normann & Sternberg, 2023). The third strand is concerned with detection and regulation, and it examines why competition law as it currently stands struggles to address this kind of behavior (Ittoo & Petit, 2017; OECD, 2017; Izzo, 2025). This review synthesizes these three strands to address a central research question: under what specific economic conditions do reinforcement-learning systems learn to collude, and what mathematical or statistical mechanisms could allow this to be detected in real markets? The sections that follow address each strand in turn before bringing them together.

2. Scope and Method

This paper is a structured literature review, meaning it does not generate new empirical data but maps and evaluates existing research. Sources were selected to cover the main strands of scholarship on algorithmic pricing and tacit collusion, with a preference for peer-reviewed articles and academic books, supplemented by policy reports where relevant. The review is interdisciplinary by necessity. Because the phenomenon involves both economic behavior and machine learning, using only one discipline would produce an incomplete picture. Economic research explains what happens in markets and why, while computer science explains how the learning algorithms function. The analysis is organized thematically rather than chronologically, which makes it easier to compare how different studies approach the same questions and to identify where they agree, diverge, or leave gaps.

3. How Algorithms Learn to Collude

The algorithms at the center of this problem are not the kind that simply follow pre-written rules. They are reinforcement-learning (RL) agents, which means they learn through experience. An RL algorithm starts without knowing what to do. It tries different actions, observes whether those actions lead to rewards or penalties, and over time adjusts its behavior to maximize its total returns. In a pricing market, the reward is profit and the action is choosing a price. What matters for understanding algorithmic tacit collusion is that this design gives the algorithm no instruction to collude. If collusion appears, it is not because someone programmed it in. It is a side effect of the algorithm optimizing for profit in a competitive environment, and this distinction has important legal implications that will come up in Section 5.

The study that established the empirical basis for this concern is Calvano et al. (2020). They ran a simulation with two Q-learning agents competing in a duopoly over thousands of repeated rounds. Q-learning is a technique where the algorithm maintains a table mapping market states to expected profits, updating it based on what it observes after each decision. The key result is that the two algorithms, without communicating, consistently ended up setting prices significantly above the competitive equilibrium and close to the monopoly level. The mechanism was a punishment strategy that the algorithms developed on their own. Whenever one agent tried to undercut the other, the rival responded with a sustained period of low prices, making the initial deviation unprofitable. Over time, both agents learned that holding the high price was the better strategy. This is important not just as an empirical result but as a conceptual one: it shows that collusion does not need communication, agreement, or any awareness of what the other firm is doing. It can emerge from simple profit maximization.

Klein (2021) tests a different setting to check how general this finding is. Where Calvano et al. used simultaneous pricing, Klein looks at sequential pricing, where firms take turns adjusting prices, which is arguably closer to how many online markets actually work. The result is similar: Q-learning algorithms still converge on collusive prices, and the sequential structure may actually make the punishment mechanism easier to execute because each agent has time to observe and react before the next adjustment. The two studies together make a reasonably strong case that the phenomenon is not an artifact of one particular model setup. That said, both rely exclusively on Q-learning, which is one of the simpler RL methods available. Commercial pricing systems use much more advanced techniques, and whether those also produce collusive behavior is largely untested. The simulations prove the concept, but there is an open question about how well the findings transfer to the real world.

4. The Conditions That Enable Collusion

Showing that algorithmic collusion can emerge in theory is one thing. Understanding when it is more or less likely in practice requires looking at market conditions. Ezrachi and Stucke (2016; 2019) offer the most cited account of this and identify two factors as particularly important. The first is reaction speed. In traditional markets, there is a delay between one firm cutting its price and rivals responding, and during that delay the firm that moved first can attract customers. That delay is the reward for competing. Algorithmic systems can eliminate it by responding to price changes in milliseconds, which removes the incentive to undercut in the first place. The second factor is price transparency. When prices are posted publicly and updated constantly online, any deviation from the high-price equilibrium is immediately visible to all other algorithms, making it very easy to trigger the kind of punishment response described by Calvano et al. (2020). What Ezrachi and Stucke's analysis adds, beyond just listing conditions, is the observation that these conditions do not sit separately from the learning mechanism. They amplify it. Speed and transparency make the punishment more effective, which reinforces the equilibrium that the algorithms have learned to sustain.

Asker et al. (2024) take a different approach and focus on the internal design of the algorithms themselves. Through a series of simulations, they show that parameters like memory length, learning rate, and discount factor all affect the prices that eventually emerge in equilibrium. This means that the way an algorithm is designed is not competitively neutral: two firms with differently calibrated algorithms in the same market can end up with very different price outcomes. There is also a practical angle here. Pricing software in many industries is sold by a small number of vendors, meaning competing firms in the same market may be running algorithms with nearly identical settings. If Asker et al. are right, that homogeneity could itself promote collusion. Normann and Sternberg (2023) raise a related issue. Most simulations model markets where all participants are algorithms, but real markets still involve human decision-makers alongside automated systems. Normann and Sternberg ran laboratory experiments mixing human participants with algorithmic agents and found that human behavior is less consistent than algorithmic behavior, which disrupts the punishment patterns that collusion depends on. Their finding does not disprove the earlier results, but it suggests those results describe something closer to a theoretical upper bound than a typical real-world outcome.

5. Detection, Legal Challenges, and Regulation

Even if we accept that algorithmic tacit collusion is real and harmful, it is not clear what can be done about it. The starting point is competition law. As the OECD (2017) explains, illegal collusion in most legal systems requires proof of an agreement or concerted practice, which requires showing communication and deliberate intent. Algorithmic tacit collusion has neither. The algorithms operate independently, never exchange information, and no human at any of the firms has decided to fix prices. There is nothing to subpoena, no internal documents to uncover, no whistleblower. The behavior produces the same economic damage as a price-fixing cartel without any of the evidence that normally makes cartels detectable. Ittoo and Petit (2017) identify a further problem. Auditing a pricing algorithm to check for anti-competitive tendencies requires technical expertise that most competition authorities do not have, and the algorithms are proprietary and constantly changing. Distinguishing between prices that are high for legitimate reasons and prices that are high because algorithms have coordinated on an equilibrium is, from the outside, very difficult.

Izzo (2025) draws the logical conclusion from this. If collusion cannot be detected after the fact, enforcement that waits for evidence will rarely find any. The traditional model of competition regulation, which investigates and punishes firms after a violation has been identified, is simply not well-suited to a world where the violation leaves no trace. Izzo proposes a shift to ex ante regulation, where firms would be required to demonstrate that their pricing systems comply with competition law before deploying them, roughly in the way that pharmaceutical companies must prove their products are safe before bringing them to market. Whether this is practically achievable is a separate debate, but the argument points to something structurally important: the legal gap and the detection gap are not independent problems. Without some method of identifying collusive behavior from observed price data, there is nothing to regulate against, regardless of what the law says. This is why, across the whole third strand of the literature, detection emerges as the central unsolved problem.

6. Synthesis: An Interlocking Problem

Looking at the three strands together, what stands out is how much they reinforce each other. The simulation literature shows that the learning mechanism exists and that it can produce collusion on its own. The conditions literature shows that certain market features, particularly speed and transparency, make the mechanism more likely to activate and harder to escape from. The detection and regulation literature shows that the behavior is almost impossible to address once it occurs, partly because existing law was not designed for it and partly because there is no agreed method for identifying it from observable data. These three problems are not parallel. They are connected. Conditions that enable collusion are also conditions that make collusion look like normal competitive pricing, which is exactly what makes it so hard to detect. And without detection, regulation cannot start. The reason algorithmic tacit collusion is such a hard problem is not any single feature of it but the way these dimensions interlock.

The literature has moved quite far in establishing that algorithmic tacit collusion is a real phenomenon. The simulation evidence is fairly robust, and the legal analysis of why it resists enforcement is well developed. What is still missing is more specific. On the conditions side, research has identified several variables that seem to matter, including reaction speed, price transparency, algorithm design, and the presence of human competitors, but these have mostly been studied one at a time. No study has yet mapped which combinations reliably produce collusion. On the detection side, the gap is even more significant. The simulation studies observe collusion directly because they have full access to the algorithms. A competition authority working from real market data has no such access, and no validated statistical method exists for distinguishing algorithmically sustained high prices from prices that simply reflect costs or demand. These gaps define what future research would need to address.

7. Conclusion

This review has examined three connected questions: how reinforcement-learning algorithms learn to collude, what conditions make this more likely, and why detecting and regulating it is so difficult. On the first question, the evidence is fairly clear. Algorithms do not need to communicate or be instructed to collude. Through repeated interaction and profit optimization, they can independently reach strategies that sustain prices above the competitive level. On the second question, the literature points to reaction speed, price transparency, algorithm design, and the presence of human competitors as the most relevant variables, though how these interact is not yet well understood. On the third question, there is broad agreement that current competition law is not suited to this behavior and that detection is the main bottleneck. A legal framework that requires proof of agreement cannot address a phenomenon that leaves no agreement behind.

These conclusions should be read with caution. The simulation evidence is based on relatively simple algorithms, and Normann and Sternberg (2023) make a fair point that real markets with human participants may behave quite differently. Two concrete research directions follow from the gaps in this review. One is to study the enabling conditions more systematically, through simulations that vary multiple factors at once rather than one at a time. The other, more pressing for policy, is to develop detection methods applicable to real market price data. Without that, competition authorities will remain unable to act even if the law eventually catches up.

Bibliography

Asker, J., Fershtman, C., & Pakes, A. (2024). The impact of artificial intelligence design on pricing. *Journal of Economics & Management Strategy, 33*(2), 276–304.

Calvano, E., Calzolari, G., Denicoló, V., & Pastorello, S. (2020). Artificial intelligence, algorithmic pricing, and collusion. *American Economic Review, 110*(10), 3267–3297.

Ezrachi, A., & Stucke, M. E. (2016). *Virtual competition: The promise and perils of the algorithm-driven economy*. Harvard University Press.

Ezrachi, A., & Stucke, M. E. (2019). Sustainable and unchallenged algorithmic tacit collusion. *Northwestern Journal of Technology and Intellectual Property, 17*(2), 217–260.

Ittoo, A., & Petit, N. (2017). Algorithmic pricing agents and tacit collusion: A technological perspective. *SSRN Electronic Journal*.

Izzo, O. (2025). *The AI cartel: Algorithmic tacit collusion and the case for ex ante regulation in EU competition law* [Unpublished master's thesis]. University of Amsterdam.

Klein, T. (2021). Autonomous algorithmic collusion: Q-learning under sequential pricing. *RAND Journal of Economics, 52*(3), 538–558.

Normann, H.-T., & Sternberg, M. (2023). Human-algorithm interaction: Algorithmic pricing in hybrid laboratory markets. *European Economic Review, 152*, 104347.

OECD. (2017). *Algorithms and collusion: Competition policy in the digital age*. OECD Publishing.

Reflection on the Process

I chose this topic because I found it genuinely interesting from the beginning. The idea that an algorithm can learn to behave like a cartel without anyone telling it to felt almost counterintuitive to me, and I wanted to understand how that was actually possible. The hardest part of writing this paper was finding the right balance between the technical side and the social and legal side. I'm doing a bachelors of science with a minor in social science (mainly economics), so the computer science and economics parts felt more natural to me. But this topic also has a big legal dimension, and I had to work harder to engage seriously with those parts. I think in the end the paper does cover both, but it took a lot of revision to make them connect properly instead of just sitting next to each other. The part I am most proud of is the synthesis, because I tried to show that the technical mechanism, the market conditions, and the detection problem are all part of the same issue, not three separate topics. I think that argument is what gives the paper a point beyond just summarizing the literature.

The feedback from my Writing Update helped a lot. My teacher asked for more historical context about how pricing worked before algorithms, so I added the part about trade publications, financial newspapers, and how the internet changed everything. He also told me to explain why over-pricing is specifically the problem and not under-pricing, which I had not thought about before, and I added that to the introduction. Another comment was to say something about who actually builds these systems, so I mentioned the software vendors that sell pricing tools to many companies at the same time. The suggestion to rethink the definition of collusion was probably the most useful one. Using the Ezrachi and Stucke distinction between messenger and autonomous collusion gave the paper a clearer starting point. I did not go as deep into the historical background as my teacher maybe wanted, but the word limit made it hard to do everything. For my Capstone, I want to focus more on the scientific and mathematical side of this topic, specifically the detection problem, because that fits better with my major and it is also the gap in the literature where I think I can actually contribute something.

AI Statement

In preparing this paper, I used two AI tools at different stages of the process. I used ChatGPT for language support. As a native Spanish speaker, I sometimes produce English phrasings that translate directly from Spanish and read unnaturally in academic prose, so I used ChatGPT to find and correct such cases and to check grammar and punctuation. I used Claude for structural and drafting support. This included help with organizing my sources into a coherent outline, drafting and revising sections of the text, and getting feedback on the clarity and strength of my arguments. In each case, I provided the research question, the source material, and my own understanding of the sources, and I reviewed, edited, and revised all output.

Sample of AI prompts used

ChatGPT prompts (language and revision support)

- "Does this sentence sound natural in academic English? I am a Spanish speaker and think I translated it too directly."
- "Correct grammar and awkward phrasing without changing my argument."
- "Which parts of this paragraph sound repetitive?"
- "Can you help me connect these two paragraphs better?"
- "Can you help make this transition between sections smoother?"

Claude prompts (structure and organization support)

- "I have sources about reinforcement learning, market conditions, and regulation. What is a logical thematic structure for a literature review?"
- "Does this paragraph actually answer the research question?"
- "How can I connect these three strands more clearly?"

- "Where do I need more analysis instead of summary?"
- "Do I explain the Calvano and Klein studies clearly enough?"
- "Do these sections connect logically or does it jump around too much?"