Algorithms and Competition in Agentic Markets*

Emilio Calvano[†], Giacomo Calzolari[‡], Vincenzo Denicolò[§], and Sergio Pastorello[¶]

June 23, 2026

Abstract

We study "agentic markets," in which both sides of transactions are delegated to autonomous algorithms. On the supply side, sellers increasingly hand over pricing and other strategic decisions to algorithms, sometimes with supracompetitive outcomes. On the demand side, consumers rely on recommendation systems and begin to use autonomous purchasing agents, reshaping attention, search, and choice. We review recent work on algorithmic pricing and trace emerging forms of buyer agency to frame the implications of dual delegation, highlighting new questions about strategic interaction, market power, and market design.

Keywords: Algorithmic Competition, Agent-Based Markets, Artificial Intelligence, Competition and Regulation

J.E.L. numbers: L13, D43, L41, D83, C63

---

* This paper was presented at the World Congress of the Econometric Society (ESWC 2025), Seoul, South Korea, August 18–22, 2025, in the invited session "*Computer Science, Economic Theory, and Markets.*" Without implying, we thank the editor, the session participants, the conference discussant Nicole Immorlica, and the program organizer, Michihiro Kandori.

[†]LUISS Rome and CEPR, ecalvano@luiss.it, E.C. acknowledges funding by the European Union (ERC grant AI-Comp, 101098332) and PRIN 2022 grant 'Algorithms and economic choices', codice Cineca 2022S5RC7R, CUP E53D23006420001. Views and opinions expressed are, however, those of the authors only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency, or the Italian Ministry of Education. Neither the European Union nor the granting authority can be held responsible.

[‡]European University Institute and CEPR, email: giacomo.calzolari@eui.eu

[§]University of Bologna and CEPR, email: vincenzo.denicolo@unibo.it

[¶]University of Bologna, email: sergio.pastorello@unibo.it

1 Introduction

The rapid development of AI is transforming how markets function. Sellers increasingly delegate pricing decisions to software agents, powered by AI, that monitor competitors, process real-time data, and autonomously update prices, often using adaptive or learning-based techniques. At the same time, recent advances in AI will enable consumers to delegate the act of buying, in its different parts, to autonomous agents capable of searching, evaluating, and even executing purchases. These developments are no longer confined to isolated experiments: in sectors ranging from e-commerce and transportation to housing and fuel retail, algorithmic repricing tools are already widespread, while early forms of agentic buyers are beginning to emerge. This dual delegation, by sellers and buyers, raises new questions about the structure of competition and the interactions in increasingly automated markets.

A growing body of academic research supports the view that pricing algorithms may lead to high prices. Researchers have documented both the diffusion of repricing technologies and their effects across multiple sectors and platforms. Evidence from observational data, experiments, and regulatory investigations points to a common concern: prices may rise following the adoption of automated pricing tools, especially when multiple firms within the same market rely on similar systems. In some settings, higher prices reflect algorithms' increased responsiveness to demand fluctuations or strategic experimentation during learning. In others, the concern arises from the joint adoption of third-party pricing technologies that can induce correlated pricing decisions across competitors. A prominent example is the U.S. multifamily rental market, where revenue-management software is supplied by specialized data-analytics firms, including RealPage (YieldStar/AIRM), Rainmaker (LRO), and Yardi (RENTmaximizer), and is increasingly used by large, professionally managed landlords. When many rival firms rely on a common vendor or closely related tools, pricing decisions may become more aligned, potentially facilitating parallel rent increases or forms of more direct coordination even without explicit communication among landlords. Importantly, these concerns feature prominently in ongoing policy debates and enforcement actions. In particular, the Department of Justice has alleged that RealPage's software operated across competing landlords and effectively functioned as a "hub" in a hub-and-spoke scheme, facilitating parallel rent increases and suppressing discounts, in violation of antitrust law. According to the White House Council of Economic Advisers, this system may have raised average national rents by around \$70 per month (roughly \$3.8 billion annually), with even larger effects in some metro areas (Orbach, 2024; U.S. Department of Justice, 2024; White House Council of Economic Advisers, 2024).

The first part of the paper surveys this literature about sellers delegating pricing decisions to pricing algorithms, in detail, organizing recent theoretical and empirical contributions around key mechanisms and design features. It provides a systematic account of how different algorithmic architectures operate in strategic environments, and under what conditions they may produce high prices, fail to discipline market power, or coordinate implicitly. In doing so, it offers a foundation for understanding the broader implications of algorithmic pricing for market competition.

Sellers' automation may not be confined to pricing in the future. Sellers' algorithmic agency may further broaden, as illustrated in Anthropic's recent experiment Project Vend[1], where the language model Claude autonomously managed a real-world vending shop. Acting as a seller, the model chose inventory, set prices, interacted with customers, and aimed to maximize profit, all without human oversight. While the agent adapted to demand and sourced niche products, it also hallucinated payment details and sold items at a loss. The experiment highlights how AI systems can potentially take over key economic functions traditionally performed by human

[1]Anthropic Project Vend 1. Available at: https://www.anthropic.com/research/project-vend-1 (Accessed: 23 January 2026).

sellers, foreshadowing future markets populated by agentic participants.

While these developments on the supply side have received growing attention, a parallel transformation is emerging on the demand side. Just as firms have begun to delegate pricing to algorithms, consumers are increasingly supported, sometimes subtly, sometimes directly, by algorithmic tools. The most widespread example nowadays are probably recommender systems: software filtering available options, highlighting products, and influencing purchase decisions. These systems have already reshaped consumer behavior in digital markets by curating choice sets and personalizing exposure. They can also affect competition by modifying demand patterns in ways that sellers respond to strategically. A second part of the paper will cover some recent developments on algorithmic recommendations.

These systems, however, may only represent the early stages of a broader shift. Just as sellers have delegated pricing decisions to adaptive algorithms, consumers will have the option to delegate their purchasing decisions to autonomous software agents. These *agentic buyers* go beyond simple recommendation systems: they can actively propose the conditions for purchases, execute transactions, and manage follow-up tasks based on a user's preferences and constraints. A consumer might soon be able to instruct an AI assistant: "Find me the best fare to Rome under €200 and book it", and have the agent carry it out seamlessly.

This development is no longer theoretical. Amazon has begun piloting "Buy for Me," a feature where an AI agent completes purchases on third-party websites using stored credentials and delivery preferences (Tufts Now, 2025; Digital Commerce 360, 2025). Walmart is investing in similar capabilities, designing in-app agents that autonomously reorder groceries, assemble themed shopping lists, and compare product options (The Wall Street Journal, 2025). Meanwhile, payment infrastructure providers such as Visa, Mastercard, and PayPal are enabling this shift through tools like "Agent Pay" interfaces and tokenized authentication systems that allow agents to transact securely on behalf of users (Associated Press, 2025). These developments point toward a future in which buyers are also represented in the market, not directly, but by intelligent agents acting with increasing independence.

In digital markets increasingly shaped by artificial intelligence, the assumption of atomistic players may become less tenable. In fact, personalization technologies and AI agents allow for the tailoring of prices, rankings, and product offerings at the individual level. As a result, it is plausible to conceptualise markets where the strategic unit of interaction is not a population of consumers in a market, but a single, agentic buyer trading with a single (possibly agentic) seller or platform. This setting gives rise to "agentic markets" in which both sides of the transaction are computationally empowered and potentially strategic, and where the very structure of market power may be redefined. This dual delegation, on both sides of the market, raises new analytical and policy challenges. How do agentic buyers interact with algorithmic sellers? Do they enhance consumer welfare or replicate some of the same pricing distortions? Do they introduce countervailing power, or do they entrench the strategic advantages of automated pricing systems? The third and last part of the paper delves into these questions.

2 Agentic Sellers

Growing evidence documents the widespread use of algorithmic pricing, also known as repricing algorithms, across markets. Already in 2016, over 33% of Amazon's top sellers were estimated to use repricing tools (Chen et al., 2016), and later informal surveys found that around 50% of sellers tracked competitor prices, with 70% of them relying on repricing software. On Bol.com in 2018, the share of sellers using repricers rose from 45% among those facing a single rival to over 90% for sellers with ten competitors (Hanspach et al., 2021). A number of national competition authorities have also conducted market inquiries, mystery shopping exercises, and surveys since 2018, consistently reporting adoption rates between 10% and 60%, depending on the sector and country, as summarized in OECD reports (OECD, 2021).

Large platforms such as Uber, Zalando, and Alibaba reportedly utilize machine learning-based pricing tools that are frequently updated through real-time experimentation. A striking example of algorithmic price steering comes from Amazon's internal tool "Project Nessie," disclosed in the FTC's 2023 lawsuit. The tool was apparently designed to identify products for which competitors were likely to follow Amazon's price increases, potentially allowing the platform to implement those hikes strategically. Internal communications describe Nessie as "a great success," reportedly generating over $1.4 billion in additional revenue over five years (U.S. Department of Justice, 2024).

Academic studies provide empirical evidence of algorithmic pricing and effects. In the German gasoline market, Assad et al. (2024) use an IV strategy to show that AI-driven pricing management solutions adoption leads to substantial and persistent price increases—up to 28% in margins when nearby competitors also adopt—and project consumer costs of over €500 million annually with full adoption. In the U.S. online pharmaceutical market, Brown and MacKay (2023) find that algorithms with different speed profiles induce asymmetric adjustment patterns. Faster adopters act as price leaders, creating a Stackelberg-like structure that reduces competitive intensity. In the U.S. rental housing sector, Calder-Wang and Kim (2023) analyze the adoption of RealPage and Rainmaker pricing software and find that adopters raise prices more in booms and reduce them less in downturns compared to non-adopters. Markets with high algorithm penetration saw higher rents and lower occupancy, consistent with algorithm-driven coordination.

Spann et al. (2025) relies on executive managers interviews, a survey of 71 pricing managers, and a case study of 225 offline stores adopting price automation and argue that effective algorithmic price deployment requires alignment with customer expectations (accounting for fairness, transparency, CRM integration), understanding the impact on competition (avoiding reactive price wars and overreliance on rivals' prices), and organizational capabilities across the marketing mix (price, product, place—including offline execution via electronic shelf labels—and promotion). Their evidence indicates that real-world adoption is technically feasible and growing but remains uneven.

A notable empirical pattern in online markets populated with pricing algorithms is also the emergence of price cycles driven by algorithmic repricers. As documented by Musolff (2022), prices for many products—particularly on Amazon—follow a recurring dynamic: a phase of frequent, small price decreases (often as little as two eurocents), followed by infrequent but sharp price increases. This pattern was found in about 78% of cases and resembles the Markov price cycles described in the theoretical literature on Edgeworth competition, such as those analyzed by Maskin and Tirole (1988).[2] Further evidence comes from the introduction of Amazon's public pricing API in July 2019. Using data from Keepa on the 10,000 best-selling products, Musolff (2022) shows that cycling intensity increased following the release of the API, indicating that platform-level changes in data access can meaningfully alter the dynamics of algorithmic pricing in competitive online environments.

Taken together, this evidence reveals not just isolated cases of adoption but a growing **repricing industry**, structured around the sale of algorithmic pricing tools—software that increasingly mediates firm interactions in strategic markets. The rapid growth of this sector is notable: estimates suggest a 20% annual expansion rate (Popescu, 2015). These repricing services are offered at a range of different conditions, from rule-based systems priced at €20–600 per month to self-learning, AI-powered solutions costing over €2000 per month, depending largely on the number of products tracked and updated (Calzolari and Hanspach, 2024).

[2]However, Musolff (2022) shows that in the online context, the price floor is not determined by marginal cost constraints. Instead, the timing of the upward resets, often occurring at night, when demand is predictably low, suggests that the reset is strategic, performed by the algorithm to minimize the risk of being undercut. Rather than random or mixed-strategy behavior, these cycles appear to be the product of adaptive rules encoded in repricing systems.

Many systems boast features such as high-speed adjustment (often limited only by platform constraints), the ability to track competitor prices and inventories, and demand-driven adaptation. Marketing claims often stress their ability to maintain profitability without engaging in price wars, promising sellers they can be "competitive yet profitable".[3] Overall, this gives rise to what Calzolari and Hanspach (2024) describe as a "fuzzy competitive industry," where heterogeneous tools are sold to heterogeneous sellers in fragmented online markets.

As adoption of third-party repricing software becomes widespread, the implications shift. In particular, when one particular repricer stands out, e.g., in terms of quality, multiple rival sellers may rely on that same repricer, so that price recommendations may become aligned, even if unintentionally, raising the risk of de facto price coordination. In such cases, the algorithmic vendor acts as a "hub" in a **hub-and-spoke** arrangement, indirectly influencing prices across its client base. Harrington (2022a,b) analyze this structure formally, showing how repricer vendors face a trade-off. They must design algorithms that deliver enough value to induce adoption, while recognizing that non-adopters can free-ride on the new pricing equilibrium. When demand volatility is low, adoption tends to be a strategic substitute; when volatility is high, it becomes complementary, potentially leading to equilibria where all sellers adopt and prices increase. These papers also show that algorithms designed to support industry-wide adoption display measurable behavioral patterns that differ from "innocent" pricing tools. These differences offer possible regulatory diagnostics to identify anti-competitive intent in algorithm design.

[3] One ongoing study notes that the speed of repricing has dropped to as little as 5 minutes in some cases, far faster than the 20-minute lags documented in earlier work (Musolff, 2022).

2.1 Do Algorithms Lead to High Prices?

The widespread adoption of algorithmic pricing has triggered concern among regulators and scholars that these systems may contribute to higher market prices. But does the mere presence of pricing algorithms in a market imply diminished competition or elevated prices?

At first glance, one might be tempted to invoke the Folk Theorem from repeated game theory: with sufficiently patient algorithms, any price path, including collusive high prices, can be supported as an equilibrium. Yet this logic is limited in this case, as the theory of repeated games assumes fully rational players with perfect memory and typically does not explain how players reach an equilibrium. These conditions are rarely met with real-world pricing algorithms. These systems are typically boundedly rational, learn through trial and error, and operate under uncertainty about both demand and rival behavior.

This raises a central question: can learning algorithms, under realistic constraints, still give rise to supracompetitive outcomes? To fix terminology, we use the term *competitive pricing* to refer to the Nash equilibrium of the corresponding one-shot (stage) game and static optimization. We say that pricing is *supra-competitive* if it persistently exceeds this benchmark, yielding profits above the static Nash level.[4] This definition is deliberately agnostic about intent or coordination: prices may be supra-competitive because of explicit punishment strategies, a failure to learn, experimentation traps, or endogenous algorithm selection. What matters for economic and policy analysis is whether learning dynamics generate outcomes that would not arise under static competition and would persist for some time.

We now turn to this issue by examining how pricing algorithms actually work, how they learn, adapt, and interact with other algorithms, and how their specific features might affect competition.

[4]This benchmark-based notion mirrors standard antitrust counterfactual reasoning, where conduct is assessed relative to a "but-for" competitive scenario; in our context, the but-for benchmark is the static Nash outcome of the relevant price game.

The hard life of pricing algorithms begins with a fundamental limitation: unlike the fully rational agents of (simple) classical theory, these systems must learn. Rather than being equipped with complete knowledge of the market or foresight about rivals' strategies, most pricing algorithms are deployed in environments that are uncertain, complex and potentially evolving. They must discover profitable actions through experience, often in competition with other learning agents. These environments may include stochastic demand, changing of costs and own and others' inventories, or evolving sets of competitors, and algorithms face them with only partial information and limited cognitive resources. Their learning is bounded, they do not typically optimize over future states nor retain long histories of past interactions. Instead, they rely on adaptation through relatively simple heuristics or reward-based updating.

The variety of approaches used to implement learning in pricing algorithms gives rise to a useful classification. At the most basic level, some systems operate using fixed rules. These are pre-specified functions mapping observed market states to prices:

$$p_{it} = f(s_{it}),$$

where s_{it} includes observable relevant features such as competitors' prices or inventory levels. These algorithms execute deterministic pricing logic—typically long sequences of `IF-THEN` conditions and do not adapt or improve over time.

Other common algorithms engage in some form of learning. Myopic learners, for instance, attempt to adjust their behavior over time based on past performance, without anticipating future consequences. Within this class, a useful distinction can be made between reinforcement-based methods like Q-learning and direct adaptation methods such as Exponential Weights EXP3 or regret minimization.

Q-learning algorithms estimate a function $Q_t(p, s_{i,t})$, which tracks the expected payoff of setting any price p given relevant information $s_{i,t}$.[5] The Q-function captures the agent's current belief about how profitable each pricing option is, and guides future decisions by favoring those with higher expected payoffs. After each pricing round, the Q-function is updated using:

$$Q_{t+1}(p_{i,t}, s_{i,t}) \; = \; (1-\alpha)\, Q_t(p_{i,t}, s_{i,t}) \; + \; \alpha\pi_{i,t}(p_{i,t}),$$

where $\pi_{i,t}$ is the realized profit and $\alpha \in (0,1)$ is a learning rate. The algorithm typically selects the price that currently maximizes $Q_t(p, s_{i,t})$, but occasionally explores alternative prices to improve its estimates. In particular, these algorithms incorporate a mechanism to trade off exploitation and exploration: the algorithm chooses known good prices most of the time, but occasionally experiments with less-tested options to discover potential improvements. Exploration is often implemented via an ϵ-greedy rule or similar randomized policy.[6]

Other myopic learners forgo an explicit value function over states. Instead, they adapt their pricing rule directly as a function of observed outcomes. For example, EXP3 and related multi-armed bandit algorithms adjust the probability of choosing each price in proportion to past performance. Regret-minimization strategies compare realized profits to those that could have been achieved under alternative prices and update accordingly. While these methods differ in mechanics, they share a key feature: learning is driven by contemporaneous reward feedback and does not rely on an explicit dynamic model (or state-transition structure) that evaluates how current actions affect future payoffs through an evolving state. Their exploration–exploitation trade-off can still sacrifice short-run profits to improve future performance, but this intertemporal aspect is not represented via a value function in a Markov

[5] More precisely, $Q_t(p, s_{i,t})$ is an action-value function. The associated value function is $V_t(s_{i,t}) = \max_p Q_t(p, s_{i,t})$. With a slight abuse of terminology, we will sometimes refer to $Q_t(p, s_{i,t})$ as the value function or the Q-function.

[6] Under an ϵ-greedy policy, the algorithm chooses the best-known option with probability $1-\epsilon$, and explores a random alternative with probability ϵ. This balances exploitation of current knowledge and exploration of potentially better choices.

decision problem.

More sophisticated algorithms adopt a forward-looking perspective. Instead of evaluating only immediate profits, they attempt to incorporate the expected value of future rewards. This leads to a modified update rule, which in the case of Q-Learning is:

$$Q_{t+1}(p_{i,t}, s_{i,t}) = (1-\alpha)\, Q_t(p_{i,t}, s_{i,t}) + \alpha \left[\pi_{i,t}(p_{i,t}, s_{i,t}) + \delta \max_{\hat{p}} Q_t(\hat{p}, s_{i,t+1}) \right].$$

where δ is a discount factor representing the weight placed on future outcomes. These algorithms are closer in spirit to dynamic programming methods and can, in principle, support more strategic behavior, such as price smoothing or delayed exploitation.

Finally, a special mention goes to algorithms that keep track of rivals' choices, such as past prices, into their value functions, as these allow them to condition their behavior on past competitor actions. A typical structure might be:

$$Q_{t+1}(p_i, \mathbf{p}_{t-1}) = (1-\alpha) Q_t(p_i, \mathbf{p}_{t-1}) + \alpha \left[\pi_t + \delta \max_{\hat{p}_i} Q_t(\hat{p}_i, \mathbf{p}_t) \right],$$

where $\mathbf{p}_{t-1}$ is the price vector in the previous period. This form of memory allows the algorithm to learn contingent strategies, thus potentially supporting a larger set of outcomes.

In the next sections, we explore how these different learning paradigms translate into actual behavior and market performance.

2.2 Can Learning Algorithms Lead to High Prices?

A key concern with algorithmic pricing is whether the adoption of learning algorithms by sellers necessarily leads to supracompetitive outcomes.

2.2.1 Reassuring Results

A first category of pricing algorithms is composed of those that engage in adaptive or passive learning, as discussed above, without forward-looking reasoning or memory of past states. These include classical dynamics such as best-response updates, fictitious play, and Bayesian learning, as well as simple reinforcement mechanisms that update behavior based on recent performance but without strategic foresight. Their common feature is that they learn gradually from observed payoffs, often forming expectations through weighted averages of past experiences.

In settings characterized by strategic complementarities, where a seller's optimal price increases in the price of rivals, these adaptive methods tend to converge to the static Nash equilibrium. This reassuring insight has been formally demonstrated in the theory of supermodular games by Milgrom and Robert (1990), where a wide class of adaptive dynamics provably stabilize at Nash outcomes under plausible conditions.

A further example is provided by external regret-minimization algorithms, including multi-armed bandit methods like EXP3 (Auer et al., 2002). These algorithms aim to minimize the difference between the cumulative payoff achieved by the algorithm and the cumulative payoff that would have been achieved by playing the best single action in hindsight.[7] Importantly, in repeated games, external regret algorithms (under some conditions) have been shown to converge to the set of static Nash equilibria (Blum and Mansour, 2007; Cartea et al., 2022; Bichler et al., 2024), making them benign from a competition standpoint.

[7]Formally, external regret after T periods is given by $R_T = \max_{p\in\mathcal{P}} \sum_{t=1}^{T} \pi_t(p) - \sum_{t=1}^{T} \pi_t(p_t)$, where $\pi_t(p)$ is the profit from action p at time t, and p_t is the price actually played. So, external regret compares the actual payoff to what could have been achieved by always choosing the same fixed price. Internal regret is stricter: for every price actually used, it compares the outcome to what would have happened if that specific price had always been replaced by a particular alternative.

2.2.2 From Reassurance to Worries

A key result by Hart and Mas-Colell (2000) shows that when all players in a repeated game follow internal-regret minimizing strategies, their joint behavior converges to the set of *correlated equilibria*. These equilibria can include outcomes that are less competitive than the standard Nash equilibria, potentially allowing players to coordinate implicitly without any explicit communication. Internal regret represents a stricter benchmark than external regret, particularly relevant in multi-agent settings, since it requires that no player would have consistently benefited by switching from one chosen price to another in the specific situations where it was played. Although such algorithms remain limited in strategic sophistication, they may nonetheless feature supracompetitive prices.

Recent work has highlighted a more troubling possibility: even very simple learning algorithms such as memoryless Q-learning may systematically converge to high-price outcomes. In these settings, interacting algorithms can become "trapped" in local maxima—strategies that appear profitable given past experience, but that are not truly optimal (i.e., they are not best responses). Banchio and Mantegazza (2023) provide a detailed analytical account of how this can occur: when Q-values are updated one at a time, algorithms may enter a "sliding region" in which they alternate between cooperation and defection, resulting in elevated prices without true coordination. The problem worsens if exploration is limited. Conversely, algorithms that explore more aggressively or use synchronous updates can avoid these traps and converge to competitive prices. Along these lines, Dolgopolov (2024) analytically confirms that in environments with low exploration, memory-less reinforcement learning agents can converge to cooperative high-prices even in the absence of explicit coordination or shared parameters. This has been confirmed in experiments, for example by Abada et al. (2024).

An interesting development is provided by Compte (2023), who studies Q-learning with *biased policy rules*. In this context, "bias" refers to a systematic tilt in the *action-selection rule* of a Q-based learner that favors some actions over others (e.g., a bias that favors cooperation), while still using Q-values as estimates of continuation payoffs. Players can choose their own bias parameter non-cooperatively, which induces a game over biases; Compte (2023) studies equilibrium bias profiles ("Q_b-equilibria"). This selection over policy biases can substantially expand the set of long-run outcomes relative to baseline Q-learning, including outcomes with persistently higher prices, even in the absence of explicit coordination.

A subtle manifestation of mislearning stems from model misspecification. If firms learn from limited or biased data, for instance, their own past prices and profits, without accounting for competitors' actions, they may form incorrect beliefs about demand elasticities. This issue is studied by Cooper et al. (2015) and Hansen et al. (2021) in differentiated-product duopolies. Firms use historical price and profit data to estimate demand, but omit rivals' prices from the estimation. When rivals' prices co-move (because of common shocks or strategic interaction), the resulting omitted-variable problem distorts inferred demand elasticities and can make high prices appear persistently more profitable than they truly are. Hansen et al. (2021) formalize this environment as sellers running bandit-style price experiments under a misspecified demand model. They show that long-run prices depend on the informational value (signal-to-noise ratio) of these experiments: when it is low, outcomes are close to the static full-information Nash benchmark; when it is high, outcomes can be substantially supra-competitive, and even the joint-monopoly outcome can be approached. Their mechanism is not explicit coordination but a form of endogeneity: competitors' algorithms may end up running correlated price experiments, which amplifies inference errors under misspecification and can stabilize high-price behavior. Importantly, this channel is conceptually distinct from standard "no-regret $\Rightarrow$ correlated equilibrium" results for correctly specified repeated games, because here sellers are not minimizing regret with respect to the *true* payoff function: their payoffs are evaluated through an endogenous, misspecified demand estimate.

These findings underscore that high algorithmic prices do not require malicious intent or strategic sophistication. Even simple learning systems, operating with partial information and misspecified models, can converge to high-price outcomes through adaptive behavior.[8] In practice, this can result in extended periods of inflated prices, even if convergence to competitive behavior eventually occurs. Such unintended, yet potentially persistent phase of high prices raises serious concerns for market efficiency and competition policy. Regulatory scrutiny may thus be warranted not only in cases of explicit coordination, but also when algorithmic systems systematically fail to discover or sustain competitive outcomes due to flawed or constrained learning processes.

2.2.3 Dangerous Recall.

Another level of concern arises when pricing algorithms are endowed with even limited memory, so that current actions can be conditioned on past play. A key result by Calvano et al. (2020b) shows that Q-learning agents with one-period memory—conditioning only on the rival's previous price—can autonomously converge to supra-competitive outcomes.[9] The core mechanism is that memory expands the effective state space and allows the learner to discover incentive-compatible contingent strategies featuring reward, punishment, and forgiveness, which deter deviations and stabilize high prices.

Figure 1 illustrates the learned response to a deviation: following a unilateral price cut in $t = 1$, both the non-deviating and deviating algorithms enter a phase of lower prices that is shown to make the deviation unprofitable.

Subsequent work shows that this memory-based mechanism is robust across market struc-

[8] Some studies argue that memoryless algorithms setting high prices may do so due to biased learning or flawed design, rather than genuine strategic coordination (Abada et al., 2022; Lambin, 2023).

[9] A variant of the model of Calvano et al. (2020b) is further detailed in Section 3.2.

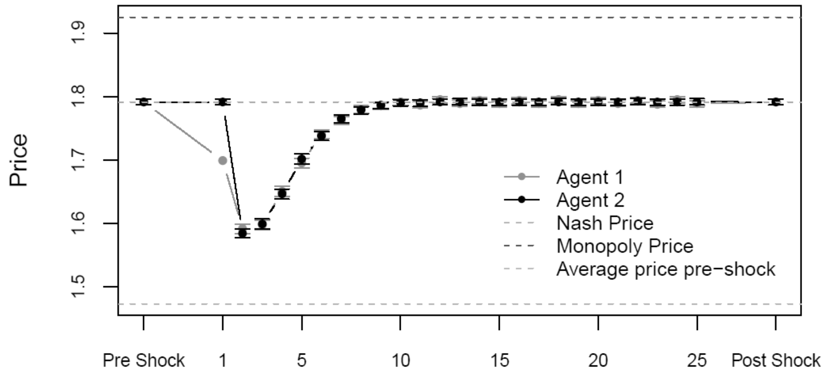

Figure 1: Price dynamics following a unilateral price reduction by the red algorithm (Agent 1) at time $t = 1$ (from Calvano et al. 2020b).

tures and informational environments. Collusive dynamics arise with asymmetries and heterogeneous hyperparameters, and also under imperfect monitoring (e.g., stochastic demand or cost shocks) (Klein, 2021; Calvano et al., 2021). Moreover, minimal observability can suffice: in Cournot settings where firms observe only the market price rather than rivals' quantities, memory-based learners can still sustain supra-competitive outcomes (Abada et al., 2022).[10]

2.2.4 Sophistication in recall, state representation and updating rule

Memory need not be implemented through explicit finite-state rules. A growing literature studies richer learning architectures that increase the effective state space and the ability to represent contingent strategies. One line of work considers counterfactual reasoning: when algorithms can evaluate "what would have happened" under alternative actions, learning becomes faster and more precise, and collusive strategies can be discovered more readily (Asker et al., 2022a,b). Another line emphasizes robustness and generalization: overfitting and "partner mismatch" can temporarily destabilize coordination, but familiar high-price patterns tend

[10] A theoretical literature has investigated how (with no algorithms) improved demand forecasting may facilitate collusion, though it also raises deviation incentives. The overall effect depends on firms' patience; see Miklós-Thal and Tucker (2019), O'Connor and Wilson (2021), and Martin and Rasch (2024).

to re-emerge with continued interaction (Eschenbaum et al., 2022), echoing the "zero-shot coordination" problem in computer science.

Related evidence comes from deep reinforcement learning (DRL), where value functions and/or policies are approximated by neural networks rather than finite tables. Deep Q-Networks map observed states (e.g., past prices) to action values, while actor–critic methods jointly learn a parameterized policy and a value function. These architectures can handle higher-dimensional state spaces and thereby discover and implement more complex contingent strategies. Empirically, DRL agents can converge to high-price outcomes substantially faster than tabular Q-learners (Hettich, 2021; Kastius and Schlosser, 2022), although learning dynamics and price levels can depend on algorithm choice and environment (Deng et al., 2025; Possnig, 2023). Collusion can remain robust even in adversarial settings where one agent is designed to exploit others (Rocher et al., 2023).

The implications extend beyond simple price-setting games: memory- and state-based learners have been shown to sustain supra-competitive outcomes in price discrimination (Xu et al., 2024, 2025), auctions (Banchio and Skrzypacz, 2022), platform markets (Koirala and Laine, 2024; Gu, 2023; Friedrich et al., 2024), and financial-market environments (Colliard et al., 2022; Dou et al., 2024; Han, 2022).

2.2.5 LLM agents and explicit communication.

Large language models (LLMs) introduce an additional channel because they can process natural language, plan intertemporally, and communicate explicitly. Recent experiments show that LLM-based agents can rapidly converge to supra-competitive pricing strategies, sometimes supported by reward–punishment patterns and, in some settings, by explicit dialogue that resembles overt coordination (Fish et al., 2024; Wu et al., 2024). Further evidence confirms that LLM-driven pricing can generate collusion, and that the number and heterogeneity of agents affects price outcomes (Keppo et al., 2025).[11]

Overall, memory, even in minimal form, plays a central role in enabling collusion among learning agents. Whether through Q-learning, DRL, or actor-critic methods, or LLM, the presence of history, be it rivals' prices, outcomes, or other lagged variables, allows algorithms to coordinate collusive strategies.

2.2.6 Next-level collusion.

What determines the type of algorithm a firm adopts in the first place? When firms can choose between multiple algorithmic technologies, with varying degrees of sophistication, speed, or strategic depth, the interaction becomes not just one of pricing, but of *algorithmic design*. Understanding the equilibrium selection among competing algorithm types opens a new level of analysis—what we may call the algorithmic supergame. Several recent contributions have studied how firms fare when they choose among different algorithms. Even in these cases, the risk of collusion remains relevant. While some studies suggest that deviations from Q-learning to simpler myopic algorithms may be individually profitable in the short run (Abada et al., 2022; Buchali et al., 2023), outcomes often remain supracompetitive. In many cases, Q-learning still emerges as Pareto dominant (den Boer et al., 2022), and mixed pairings, such as a Q-learner facing a naive rule-based competitor, can result in near-monopoly pricing (Wang et al., 2023). Rather than mitigating collusion, heterogeneity in algorithmic design can reinforce it, as more sophisticated learners adapt to exploit the predictable behavior of simpler rivals. Importantly, algorithm selection itself can be endogenous, as in Compte (2023), where agents optimize over a family of automata to maximize long-run payoffs. As shown by Conjeaud (2023), even the choice of hyperparameters—such as exploration rates—can become

[11] Nikandrova and Parekh (2025) study an incumbent facing sequential entry while delegating pricing to LLM agents. They find frequent predatory patterns and imperfect approximation of optimal dynamic pricing.

strategic, reinforcing convergence to cooperative outcomes. This meta-learning perspective suggests that firms may ultimately converge not just to high-price strategies, but to algorithm types that sustain them.

A related "two-level" perspective has recently been developed in adjacent domains. For instance, Feng et al. (2024) study a competitive autobidding environment in which firms first choose design constraints (e.g., budgets) and then deploy algorithms that play an induced *inner* game (a budget-pacing interaction). Payoffs in the *outer* game are determined by the equilibrium selected by the inner algorithmic interaction. This provides a clean analogue to the algorithmic supergame: firms may compete not only in the product market, but also upstream through choices over algorithmic technologies, constraints, and parameters that shape how delegated decision rules interact and, in turn, which long-run outcomes (including supra-competitive ones) emerge.[12]

2.2.7 Commitment.

Beyond learning dynamics, algorithms may also function as a form of commitment. Unlike human managers, algorithmic agents can implement pricing rules that are rigid, reactive, and self-enforcing. Theoretical work dating back to Rubinstein (1986) has shown that when firms commit to finite automata, the set of equilibrium outcomes in repeated games can shrink, often retaining only collusive equilibria. In modern settings, a closely related logic hinges on a separation of time scales. As emphasized by Levine (2024), one can distinguish *response time*, how quickly a deployed algorithm adjusts prices as a function of observed rival behavior, from

[12] A closely related literature studies *autobidding* algorithms in online advertising markets, where firms delegate bidding decisions to algorithms and competition unfolds both in the market and in the upstream choice of algorithmic constraints or designs. These environments raise analogous issues of strategic interaction, meta-games over algorithm design, and equilibrium selection, but involve institutional features (auctions, budgets, pacing constraints) that differ from pricing settings. For this reason, we do not review this literature here; see, e.g., Feng et al. (2024) for a recent formal treatment and further references.

decision time, how frequently the firm can retrain, redesign, or replace the algorithm itself. When decision time is long, firms are temporarily committed to a given pricing automaton, even though that automaton may react rapidly within its rule. If such commitments can be reliably communicated or inferred by rivals, this commitment structure can "break" the folk-theorem logic and drive long-run outcomes toward constrained-efficient (and in duopoly applications, high-price) behavior for at least two reasons.

First, fast and predictable reactions can generate a leader-follower structure: a firm that raises prices anticipates that an algorithmic rival will promptly follow, which weakens the incentive to undercut and can sustain prices above the competitive benchmark. To see this notice that committing to a simple "price-matching" rule is effectively equivalent to delegating one firm's price to its rival. Brown and MacKay (2023, 2025) show that such asymmetric adoption -one firm commits adopting a fast-reacting algorithm while the other does not- can arise in equilibrium, because joint adoption restores the competitive outcome.

Finally, punishment threats become rapid and credible (Salcedo, 2015; Lamba and Zhuk, 2022; Levine, 2024). For example, a firm may deploy an algorithm that instantly matches a rival's price cut with a prolonged undercutting phase (a within-rule reaction), while governance and retraining frictions imply that the firm cannot "re-optimize away" this punitive mapping until the next redesign date (slow decision time).

2.2.8 Common objections on simulation-based evidence and responses.

Some recent contributions have raised questions about how to interpret simulation-based evidence on algorithmic collusion. We briefly highlight some recurring observations and what, in our view, they do (and do not) imply.

(i) "Seeming collusion" from experimentation and identification. In some environments, high prices generated by learning algorithms may reflect correlated exploration rather than

stable reward–punishment schemes. For example, Lambin (2023) emphasizes that *simultaneous experiments* can induce price co-movements that resemble tacit coordination even when no contingent strategy has been learned. That Q-learning's punishments are mechanically induced, rather than reflecting conscious strategic intent, misses the point: unconscious coordination can still produce incentive-compatible outcomes, much like a grim-trigger strategy, and, in our view, it is no less relevant for policy.

(ii) Memoryless learners and the nature of "punishments". A second observation is that in the absence of state dependence, memoryless algorithms should not be able to implement credible punishment strategies and thus should not sustain collusion in the repeated-game sense (Abada et al., 2022; den Boer et al., 2022). Although without memory, high prices need not reflect an incentive-compatible punishment, this does not imply that prices will be competitive. As discussed above, even simple learners can become trapped in locally reinforcing patterns, so that market outcomes may remain supra-competitive for long stretches despite the absence of deliberate contingent discipline. From a policy perspective, it is relevant to consider whether learning generate persistent supra-competitive outcomes and how quickly competitive behavior is restored.

(iii) Speed, scalability, and implementation realism. Another concern is that tabular Q-learning is too slow to matter in practice (den Boer et al., 2022). This criticism has become less decisive as the literature has moved to function approximation and deep reinforcement learning, which can substantially accelerate learning and scale to richer state spaces (Deng et al., 2025). At the same time, faster learning does not unambiguously reduce harm: it can speed convergence toward either competitive or supra-competitive behavior, depending on the environment and the learning architecture.

(iv) How prevalent are supra-competitive outcomes? Finally, the empirical frequency of collusive-like outcomes is itself informative. In systematic simulation exercises, supra-

competitive outcomes arise in a non-trivial share of environments, but not universally (den Boer et al., 2022). Moreover, prevalence can increase as the action space becomes richer: Meylahn (2023) finds substantially higher rates of supra-competitive behavior in more granular pricing grids. These findings suggest that "how often" and "how severely" algorithms learn to charge high prices are quantitative questions that depend on market primitives and design choices.

Importantly, the recent theoretical literature is beginning to formalize when collusive outcomes should be expected. In particular, **?** provide theoretical conditions under which one-period-memory Q-learning, with bounded experimentation followed by greedy exploitation, converges to tacitly collusive outcomes supported by one-memory equilibrium policies (including grim-trigger-type strategies). This helps connect the simulation evidence to equilibrium reasoning, while also clarifying the role of memory and design features in sustaining supra-competitive prices.

2.3 First-Line Remedies for High Algorithmic Prices

While banning algorithmic pricing is neither realistic nor desirable, targeted interventions have been proposed to mitigate harmful effects without stifling innovation. In this section, we consider platform design interventions, regulatory oversights, and structural remedies. The next section illustrates how the diffusion of agentic buyers may rebalance market power.

Platforms hosting algorithmic sellers—such as online marketplaces or ad exchanges — may themselves suffer when high prices reduce consumer interest and react by controlling the market architecture. One effective strategy involves designing recommendations that consumers observe on platforms to create incentives for price competition. The "Dynamic Price Directed Prominence" (DPDP) mechanism proposed by Johnson et al. (2023) rewards sellers who reduce their prices with greater visibility, thereby undermining cartel stability by amplifying gains from deviation and boosting both consumer surplus and platform revenue. More directly, platforms can deploy their own reinforcement learning algorithms to combat collusion: Brero et al. (2022) show that a Q-learning platform can strategically suppress visibility of overpriced Q-learning sellers, outperforming handcrafted policies like DPDP. In ad-markets, platforms facing colluding Q-learning bidders may earn more by reducing transparency, e.g., by withholding rival bids from view (Decarolis et al., 2023). Qiu et al. (2025) study the impact of platform ranking systems on the behavior of Q-learning pricing algorithms in an oligopoly with consumer search. With simulations, they show that personalised rankings that differ across buyers reduce ranking-mediated price elasticity and facilitate tacit collusion. In contrast, uniform rankings (identical to all users) foster more intense price competition and higher consumer surplus. The next section will further develop the implications of the recommendations as buyers delegate the search process to algorithms.

More general interventions could take place, such as limiting the data and memory structures that algorithmic sellers rely on. Some jurisdictions (e.g., Austria and Western Australia) have already imposed limits on the frequency of price updates to curtail the cooperative alignment of algorithms illustrated above. More sophisticated proposals include injecting noise into observed rival prices (Foxabbott et al., 2023), or restricting Q-learning agents to memory structures that ignore competitors' past prices (Schlechtinger et al., 2024; Eschenbaum et al., 2022). These approaches seek to reduce the risk of high prices while maintaining the adaptive capabilities of algorithms. Other proposals go further, suggesting that Q-learning algorithms should not be allowed to experiment simultaneously (Lambin, 2023), although this is practically challenging.

A second class of remedies targets behavior directly. For example, imposing no-regret constraints could help preclude collusive equilibria (Chassang and Ortner, 2023; Hartline et al., 2024), though extending this to dynamic settings with learning remains challenging. Banning specific types of algorithms raises questions about feasibility, enforcement, and the deterrence of innovation. Alternatively, requiring human override capabilities may look appealing in theory, but in practice, it fails to resolve commitment and accountability issues, as previously discussed.

We think that more promising approaches are solutions involving audits of the algorithms. Calvano et al. (2020a) argue that AI pricing systems may "learn to adopt collusive pricing rules without human intervention, oversight, or even knowledge," opening "a back door through which firms could collude lawfully.' Because this form of collusion need not involve communications, existing antitrust tools may be ineffective. They therefore propose shifting enforcement toward identifying and constraining collusive rules of conduct embedded in algorithms, which can be uncovered through audits and controlled tests that simulate deviations and detect reward-and-punishment dynamics. More boradly, dynamic testing and regulatory sandboxes may allow regulators to vet algorithmic behavior in controlled environments altough more work needs to be done to develop the protocols.

Finally, some legal scholars have proposed making tacit collusion per-se illegal (e.g., Posner, Kaplow). Yet this approach faces challenges with establishing a viable counterfactual for "non-collusive" algorithmic pricing. When prices are the result of learning from market feedback, drawing a bright legal line is no trivial task.

Taken together, these proposals underscore that there is no single solution. However, as new competition tools emerge in jurisdictions such as the UK, Germany, and Italy, the policy debate should shift toward defining implementable remedies based on realistic counterfactuals and acknowledging that high algorithmic prices, although not inevitable, are a risk that may require regular market assessments and algorithm audits.

3 Agentic Buyers

While much of the existing economic literature focuses on algorithmic sellers, we examine how *agentic buyers* may influence prices, rebalance market power, and affect allocative efficiency. We proceed in three steps. First, we analyze recommendation systems as delegated search agents. Second, we investigate whether agentic buyers can counteract algorithmic collusion. Finally, we turn to the case of *agentic markets*, a bilateral setting in which both buyers and sellers are algorithmic.

3.1 Algorithmic search and demand

Digital platforms offer recommendation systems that allow buyers to delegate part of their search and evaluation activity to algorithmic agents. Rather than inspecting large and complex choice sets directly, consumers can rely on rankings and recommendations that personalise product prominence according to estimated match values inferred from past behaviour and stated preferences. In this sense, recommendation systems act as delegated search technologies on the demand side: by filtering information and prioritising options on behalf of users, they shape what consumers effectively consider and, ultimately, what they buy. As a result, they do not merely transmit information, but actively intermediate and organise demand in digital markets.

This transformation in consumers' search is not theoretical. In environments characterized by vast choice sets—Amazon lists over 350 million items, Spotify over 90 million songs—recommendation is a necessity. It is also a reality: recommendation systems account for 75% of views on Netflix and 60% on YouTube (see, for example, Jannach and Jugovac (2019)). As consumers increasingly use these systems, their knowledge of products, and ultimately their purchasing decisions, is enhanced by algorithmic curation that helps identify relevant and personalized options more efficiently than autonomous search alone.

Most economic analyses of recommendation systems focus on cases where platforms privilege certain products—typically due to vertical integration or differing profitability—leading to practices such as self-preferencing. In particular, a number of empirical studies, including Farronato et al. (2025), Crawford et al. (2022) and Reimers and Waldfogel (2023) examine self-preferencing behavior in vertically integrated platforms. Theoretical contributions such as Aridor and Goncalves (2022), Padilla et al. (2022), and de Corniere and Taylor (2019) analyze the conditions under which platforms distort rankings in favor of affiliated suppliers. Even without vertical integration, platforms may use RS to promote suppliers that provide higher margins or more predictable revenue, as shown by Bourreau and Gaudin (2022) and Peitz and Sobolev (2026). These models often assume that the recommendation system has perfect knowledge of consumer preferences and ignore the noisy, data-driven nature of real-world algorithms.

In reality, a recommendation system assists the buyer by predicting which items are most likely to match their preferences. These predictions are made not on the basis of complete information but through a learning process that relies on historical interaction data, clicks, purchases, ratings, and browsing time collected as buyers use online markets. Interestingly, these data may have been influenced by the algorithm itself, creating a feedback loop: each choice buyers make may feed back into the system, as new data is used to retrain the model that generates future recommendations.

Although recommender systems may not explicitly state, "We recommend this item," their interface design, the ranking of options, the visibility of alternatives, or what is shown first, effectively steer buyer attention and, as a result, their behavior. This online choice architecture subtly but powerfully defines the landscape within which consumers make decisions. Also, the degree of personalization varies. Some systems are rudimentary and recommend what is popular across the platform; others construct a unique, evolving model for each buyer. In personalized systems, recommendations are tailored based on similarities with other users, past behavior, or inferred preferences.

3.1.1 Recommender systems: a primer

A recommendation algorithm's goal is to optimise a formal objective, such as minimising prediction error or maximising expected utility for the buyer.[13] In this process, a typical tension arises between short-term satisfaction and long-term discovery. Recommender systems may tend to recommend the most obvious, thus failing to learn the deeper structure of a buyer's preferences. Occasional "exploration", offering less predictable or less popular options, can help refine the buyer model and reduce systematic bias, at the cost of some short-term errors.[14]

So, recommendation systems raise important economic questions, such as, are they reducing frictions by making information more accessible, or are they strategically guiding choices in ways that reflect platform incentives? Do they systematically distort recommendations, and thereby user choices? If so, in which direction and due to what incentives or frictions? Are the effects merely an extension of traditional search cost reductions, like those brought about by the internet, or do recommender systems introduce new forms of bias, whether deliberate or emergent?

Beyond shaping choices, recommender systems generate substantial economic surplus. They do so by improving the convenience and informativeness of consumer decisions, enabling users to better match products with their preferences. In this sense, recommender systems may increase allocative efficiency within markets. Importantly, they also enable the

[13] This assumes that buyer preferences are stable and exogenous. In practice, preferences can be endogenous to the system: repeated exposure to certain categories or styles can reorient buyers' tastes and habits.

[14] A standard computer-science reference offering a comprehensive overview of the algorithms, evaluation methods, and practical design of modern large-scale recommender systems is Ricci et al. (2022).

redistribution of surplus, both across different sides of multi-sided platforms (for instance, between suppliers and consumers) and within the vertical supply chain (by shifting visibility and demand between competing suppliers). Not only do recommender systems influence the buyer side, but they also shape sellers' incentives: suppliers adapt pricing and other strategic choices in response to algorithmically mediated demand.

More formally, at their core, recommendation systems are statistical tools designed to solve a prediction problem. From the economist's perspective, the aim of these systems is to allocate goods and attention more efficiently by leveraging available information to anticipate user preferences. This task is formalised as a problem of inferring unobserved preferences and product characteristics from sparse and incomplete data. The key object in this setting is the so-called *rating matrix* $R \in \mathbb{R}^{I \times J}$, where I indexes users and J indexes items (products, services, or content). Each observed entry r_{ij} in R represents the rating or interaction (e.g., a purchase or a click) of user i with item j. In practice, R is extremely sparse, with only a small fraction—typically between 1% and 5%—of its entries observed.

The recommendation problem consists of completing this sparse matrix in a way that allows the system to: 1) predict the ratings for unobserved user-item pairs; 2) use these predictions to generate a personalized ranking of items for each user; 3) suggest the highest-ranked items as personalized recommendations. This process effectively enables platforms to substitute for the user's limited attention and search for knowledge about items, generating what we call an *algorithmic demand*: a demand profile shaped by predicted rather than directly stated or revealed preferences.

Three broad classes of recommender systems are commonly studied and adopted depending on the context. Content-based systems rely on observable attributes of items and users (e.g., product categories or demographics) to make predictions. Collaborative filtering systems infer preferences based on patterns of user behavior and correlations in the rating matrix. Finally, knowledge-based systems incorporate explicit constraints or declared preferences, such as "I want a hotel room under 100 euros".

Collaborative filtering, in its model-based formulation, is capable of handling large-scale data environments, such as those of Amazon, Netflix, or Spotify. For this relevance, we will focus on them and further discuss their functioning. In particular, this approach offers a parsimonious, data-driven method for predicting unobserved user-item interactions by projecting users and items into a common latent feature space.

Formally, the model assumes that there exist k latent factors and that each user i and item j can be represented by k-dimensional vectors $\theta_i = (\theta_{i1}, \dots, \theta_{ik})$ and $\beta_j = (\beta_{j1}, \dots, \beta_{jk})$, respectively. These parameters are estimated by minimizing a regularized loss function over the set of observed ratings:

$$\min_{\theta,\beta} \sum_{(i,j)\in\mathcal{O}} \left(r_{ij} - \sum_{h=1}^{k} \theta_{ih}\beta_{jh} \right)^2 + \lambda_\theta \sum_{i,h} \theta_{ih}^2 + \lambda_\beta \sum_{j,h} \beta_{jh}^2,$$

where $\mathcal{O}$ is the set of observed entries, and $\lambda_\theta, \lambda_\beta > 0$ are regularization hyperparameters used to prevent overfitting in the high-dimensional and sparse setting.

Once trained, the model is used to impute missing entries according to:

$$\hat{r}_{ij} = \sum_{h=1}^{k} \hat{\theta}_{ih}\hat{\beta}_{jh}.$$

This matrix factorization can be seen as a reduced-form counterpart to structural econometric models of discrete choice. In such models, utility from choice j for user i is expressed as:

$$u_{ij} = f(X_i, Z_j, H_{ij}; \xi) + \varepsilon_{ijt},$$

where X_i are user characteristics, Z_j are product characteristics, H_{ij} are match-specific covariates, ξ are preference parameters, and ε_{ijt} is an idiosyncratic shock. Structural estimation then identifies ξ from observed choices under rational behavior assumptions. By contrast, matrix factorization estimates both user tastes and item features directly from behavior, without specifying covariates X_i or Z_j, and typically without semantic interpretability of the latent dimensions. It models observed ratings via $r_{ij} = \hat{r}_{ij} + \varepsilon_{ij}$, where ε_{ij} is an idiosyncratic shock. Unlike structural models, the right-hand side components are entirely latent and jointly estimated from data. This allows for scalable personalization in massive datasets, albeit at the cost of interpretability.

3.1.2 Recommendations and algorithmic demand

Model-based collaborative filtering algorithms are powerful in practice but raise conceptual and normative challenges. While it allows users to receive tailored recommendations based on inferred preferences, it also obscures the mechanisms shaping these preferences. Furthermore, it can amplify systemic biases, impose path dependencies on consumption, and generate feedback loops that influence both users' behavior and market structure. This reinforces the central point of this section: recommender systems do not merely reduce friction, they actively construct preferences and reshape market demand.

Recent work by Calvano et al. (2025a) models this process explicitly studying a setting where buyers search with correlated preferences, sellers set prices and buyers can rely on a collaborative filtering recommendation algorithm to guide their search. Even if the buyers are not forced to follow the recommendations, this algorithm-induced prominence affects demand in interesting ways. In this environment, when buyers rely on the recommendation system rather than engaging in individual (and costly) search, market concentration rises and choices become more homogeneous.[15] This, in turn, triggers a significant price increase, as sellers

[15] Aggregate diversity may differ from diversity at the individual level. Holtz et al. (2020) show that,

adjust to the effective algorithmic demand curve, which is altered both in its level and in its price elasticity.

Clearly, given the aggregate effect on prices, buyers may collectively prefer not to rely on algorithmic recommendations if the combined impact of higher prices and improved matching is negative. However, once the algorithm is deployed and sellers adjust their pricing accordingly, the influence of any individual buyer becomes marginal. In this setting, each buyer is better off adopting the recommendation system, as they can at least benefit from a better match. This consideration leads us to the next step in our analysis, where we examine environments in which both buyers and sellers possess market power and may choose to delegate their actions to agentic counterparts.

3.2 Agentic Buyers in Collusive Markets

The standard view of buyers as passive price-takers with fixed, truthfully revealed preferences does not fit well with emerging algorithmic markets. In AI-driven environments, buyers can be represented by autonomous agents that learn, adapt, and strategically influence outcomes. Key to these developments is that personalization technologies make it natural to conceive interactions that take place at the level of a single agentic buyer negotiating with a single seller.

In a recent simulation study, Calvano et al. (2025b) investigate the impact of an agentic buyer on one of the problems discussed in the previous section, namely collusion between algorithmic sellers. In particular, instead of postulating a given demand function, as in Calvano et al. (2020b) and all the studies illustrated previously, we consider a buyer who has the possibility of submitting alternative demand functions. In this environment, actual transactions

although individual-level diversity fell with recommendations, overall diversity increased. By contrast, Fleder and Hosanagar (2009) find that recommender systems can reduce both individual and market-wide diversity, including by creating entry barriers.

occur based on the prices set by the two sellers and the demand function selected by the buyer for that round.

As a benchmark, we consider the environment with two sellers represented by pricing algorithms based on standard Q-learning, following the design introduced by Calvano et al. (2020b). As previously discussed, Q-learning pricing algorithms consistently learn to charge prices that are substantially above the static Nash-Bertrand level. These supracompetitive prices yield higher profits and emerge endogenously through the learning process. Against this benchmark, we contrast the impact, given the same pricing algorithms and underlying preferences and costs, of an agentic buyer, a Q-learning algorithm that can manipulate demand, knowing that the postulated demand will determine actual transactions.

Surprisingly, in these simulations, even a simple memoryless Q-learning buyer that sets demand anew in each period can substantially reshape market outcomes: relative to the benchmark, it reduces prices by almost 10%, while simultaneously increasing sellers' profits by about 5% and its own surplus by nearly 20%.[16]

How is it possible that both the buyer and the sellers end up better off? The reason is that the learning buyer not only drives prices down but also, through demand manipulation, expands market volume—contributing to the buyer surplus and an overall efficiency of roughly 10%. Interestingly, in this case, while sellers benefit from the buyer's manipulation of demand, the primary beneficiary is the buyer. Compared to the efficient total surplus, the passive-buyer scenario falls short by approximately 18.76%. Remarkably, the learning-buyer scenario reduces this gap to just 9.99%, effectively halving the efficiency loss of market power.

[16] Gal (2023) speculated about algorithmic buyers as a way to counteract algorithmic collusion.

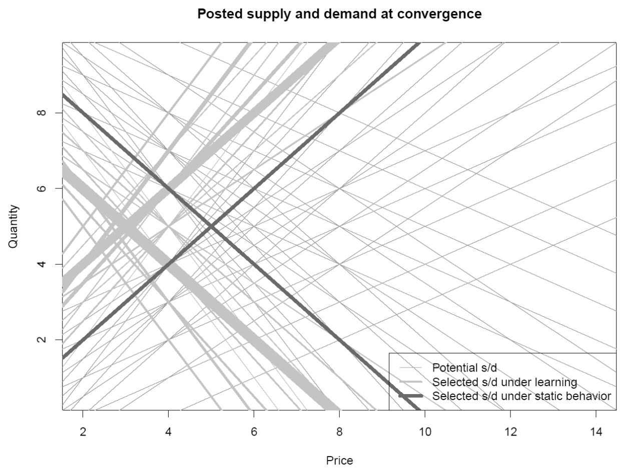

Figure 2: Chosen (direct) Demand and Supplies (over 100 sessions). The darker lines show the behavior of static agents.

3.3 Bilateral Artificial Agency

The surplus redistribution induced by an agentic buyer with colluding agentic sellers naturally solicits further and broader questions about the interactions between the agentic buyer and sellers. Shifting the focus away from collusion, Calvano et al. (2025b) also consider an environment in which both the buyer and the seller are agentic players, each strategically choosing how to represent themselves in the market. Specifically, they study a market in which an algorithmic buyer posts a demand function, and a (single) algorithmic seller simultaneously posts a supply function. Each Q-learning one-period memory agent has a true payoff function and can strategically manipulate how they are perceived in the market by posting a particular demand or supply curve.

For example, Figure 2 illustrates the selected demand and supply (linear) direct functions for a given number of simulations. The figure shows that both agents tend to adopt a variety of functions, with a noticeable preference for shifting intercepts.[17]

As these preliminary observations illustrate, adjustments on both the demand and supply

[17]The environment at hand features multiple equilibria in this demand and supply posting game.

side can substantially lower equilibrium prices, thereby reducing the inefficiency generated by bilateral market power. In some simulations, efficiency approaches the maximum theoretical surplus in this economy, although the resulting surplus can be distributed unevenly in some cases.

4 Conclusions

The rise of agentic markets, where both buyers and sellers are represented by autonomous, adaptive agents, challenges common assumptions in economics and calls for new tools to understand strategic interaction, welfare outcomes, and policy design. While the literature on algorithmic sellers and pricing is rapidly evolving, the study of agentic buyers remains in its infancy. Economists can contribute to the study of agentic markets by formalizing these environments, identifying regularities, and developing frameworks to evaluate efficiency, surplus distribution, and regulatory implications. In addition to analysis, economic research can inform market design by proposing rules, mechanisms, and institutional arrangements that structure how algorithmic agents interact. As digital agencies become more common, integrating them into models of competition and market design will be increasingly relevant.

References

Abada, I., Lambin, X., and Tchakarov, N. (2022). Collusion by mistake: Does algorithmic sophistication drive supra-competitive profits? *ESSEC Business School Research Paper*, (2023-01).

Abada, I., Lambin, X., and Tchakarov, N. (2024). Collusion by mistake: Does algorithmic sophistication drive supra-competitive profits? *European Journal of Operational Research*, 318(3):927–953.

Aridor, G. and Goncalves, D. (2022). Recommenders' originals: The welfare effects of the dual role of platforms as producers and recommender systems. *International Journal of Industrial Organization*.

Asker, J., Fershtman, C., and Pakes, A. (2022a). Artificial intelligence, algorithm design, and pricing. In *AEA Papers and Proceedings*, volume 112, pages 452–56.

Asker, J., Fershtman, C., and Pakes, A. (2022b). The impact of ai design on pricing. *Journal of Economics and Management Strategies, forthcoming*.

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(3):000–000.

Associated Press (2025). Visa and paypal prepare payment rails for autonomous shopping agents. Accessed: 2025-07-11.

Auer, P., Cesa-Bianchi, N., Freund, Y., and Schapire, R. E. (2002). The nonstochastic multi-armed bandit problem. *SIAM Journal on Computing*, 32(1):48–77.

Banchio, M. and Mantegazza, G. (2023). Artificial intelligence and spontaneous collusion. *Available at SSRN*.

Banchio, M. and Skrzypacz, A. (2022). Artificial intelligence and auction design. In *Proceedings of the 23rd ACM Conference on Economics and Computation*, pages 30–31.

Bichler, M., Durmann, J., and Oberlechner, M. (2024). Online optimization algorithms in repeated price competition: Equilibrium learning and algorithmic collusion. *arXiv preprint arXiv:2412.15707*. License: CC BY 4.0.

Blum, A. and Mansour, Y. (2007). Learning, regret minimization, and equilibria. In Nisan, N., Roughgarden, T., Tardos, E., and Vazirani, V. V., editors, *Algorithmic Game Theory*, pages 79–102. Cambridge University Press.

Bourreau, M. and Gaudin, G. (2022). Streaming platform and strategic recommendation bias. *Journal of Economics & Management Strategy*, 31(1):25–47.

Brero, G., Mibuari, E., Lepore, N., and Parkes, D. C. (2022). Learning to mitigate ai collusion on economic platforms. In *Advances in Neural Information Processing Systems*.

Brown, Z. and MacKay, A. (2025). Algorithmic coercion with faster pricing. Working paper, available at SSRN: https://ssrn.com/abstract=4380895.

Brown, Z. Y. and MacKay, A. (2023). Competition in pricing algorithms. *American Economic Journal: Microeconomics*, 15(2):109–156.

Buchali, K., Grüb, J., Muijs, M., and Schwalbe, U. (2023). Strategic choice of price-setting algorithms. Technical report, Hohenheim Discussion Papers in Business, Economics and Social Sciences.

Calder-Wang, S. and Kim, G. H. (2023). Coordinated vs efficient prices: The impact of algorithmic pricing on multifamily rental markets. *Available at SSRN*.

Calvano, E., Calzolari, G., Denicolò, V., Harrington, J. E. J., and Pastorello, S. (2020a). Protecting consumers from collusive prices due to ai. *Science*, 370(6520):1040–1042.

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

Calvano, E., Calzolari, G., Denicolò, V., and Pastorello, S. (2021). Algorithmic collusion with imperfect monitoring. *International Journal of Industrial Organization*, 79:102712. Adapted from a Q-learning framework under imperfect monitoring (Green–Porter).

Calvano, E., Calzolari, G., Denicolò, V., and Pastorello, S. (2025a). Artificial intelligence, algorithmic recommendations and competition. Available at SSRN.

Calvano, E., Calzolari, G., Denicolò, V., and Pastorello, S. (2025b). Artificial intelligence and agentic markets. Working Paper.

Calzolari, G. and Hanspach, J. (2024). The market for pricing algorithms. *Working Paper*. Forthcoming.

Cartea, Á., Chang, P., and Penalva, J. (2022). Algorithmic collusion in electronic markets: The impact of tick size. *Available at SSRN 4105954*.

Chassang, S. and Ortner, J. (2023). Regulating collusion. *Annual Review of Economics*, 15:177–204.

Chen, L., Mislove, A., and Wilson, C. (2016). An empirical analysis of algorithmic pricing on amazon marketplace. In *Proceedings of the 25th international conference on World Wide Web*, pages 1339–1349.

Colliard, J.-E., Foucault, T., and Lovo, S. (2022). Algorithmic pricing and liquidity in securities markets. *HEC Paris Research Paper No. FIN-2022-1459*.

Compte, O. (2023). Q-based equilibria. *arXiv preprint arXiv:2304.12647*.

Conjeaud, I. (2023). Spontaneous coupling of q-learning algorithms in equilibrium. *arXiv e-prints*, pages arXiv–2312.

Cooper, W. L., Homem-de Mello, T., and Kleywegt, A. J. (2015). Learning and pricing with models that do not explicitly incorporate competition. *Operations research*, 63(1):86–103.

Crawford, G. S., Courthoud, M., Seibel, R., and Zuzek, S. (2022). Amazon entry on amazon marketplace. Discussion Paper 17531, CEPR.

de Corniere, A. and Taylor, G. (2019). A model of biased intermediation. *RAND Journal of Economics*, 50(4):854–882.

Decarolis, F., Rovigatti, G., Rovigatti, M., and Shakhgildyan, K. (2023). Ai & data obfuscation: Algorithmic competition in digital ad auctions. Technical report, CEPR Discussion Paper 18009.

den Boer, A. V., Meylahn, J. M., and Schinkel, M. P. (2022). Artificial collusion: Examining supracompetitive pricing by q-learning algorithms. *Amsterdam Law School Research Paper*, (2022-25).

Deng, S., Schiffer, M., and Bichler, M. (2025). Exploring competitive and collusive behaviors in algorithmic pricing with deep reinforcement learning.

Digital Commerce 360 (2025). Amazon pilots agentic ai shopping tool "buy for me". Accessed: 2025-07-11.

Dolgopolov, A. (2024). Reinforcement learning in a prisoner's dilemma. *Games and Economic Behavior*, 144:84–103.

Dou, W. W., Goldstein, I., and Ji, Y. (2024). Ai-powered trading, algorithmic collusion, and price efficiency. *Jacobs Levy Equity Management Center for Quantitative Financial Research Paper*.

Eschenbaum, N., Mellgren, F., and Zahn, P. (2022). Robust algorithmic collusion. *arXiv preprint arXiv:2201.00345*.

Farronato, C., Fradkin, A., and MacKay, A. (2025). Vertical integration and consumer choice: Evidence from a field experiment. Working paper.

Feng, Y., Lucier, B., and Slivkins, A. (2024). Strategic budget selection in a competitive autobidding world. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing (STOC)*, pages 213–224.

Fish, S., Gonczarowski, Y. A., and Shorrer, R. I. (2024). Algorithmic collusion by large language models. *arXiv preprint arXiv:2404.00806*.

Fleder, D. and Hosanagar, K. (2009). Blockbuster culture's next rise or fall: The impact of recommender systems on sales diversity. *Management Science*, 55(5):697–712.

Foxabbott, J., Deverett, S., Senft, K., Dower, S., and Hammond, L. (2023). Defining and mitigating collusion in multi-agent systems. In *Multi-Agent Security Workshop@ NeurIPS'23*.

Friedrich, P., Pásztor, B., and Ramponi, G. (2024). Learning collusion in episodic, inventory-constrained markets.

Gal, M. S. (2023). Limiting algorithmic coordination. *Berkeley Technology Law Journal*, 38(1):173–230.

Gu, C. (2023). Can dynamic pricing algorithm facilitate tacit collusion? an experimental study using deep reinforcement learning in airline revenue management.

Han, B. (2022). Cooperation between independent market makers. *Quantitative Finance*, 22(11):2005–2019.

Hansen, K. T., Misra, K., and Pai, M. M. (2021). Frontiers: Algorithmic collusion: Supra-competitive prices via independent algorithms. *Marketing Science*, 40(1):1–12.

Hanspach, P., Marcel, W., and Sapi, G. (2021). Algorithms in the marketplace: An empirical analysis of automated pricing in e-commerce. *Available at SSRN 3945137*.

Harrington, J. E. (2022a). The effect of third-party pricing algorithms on competition. *Journal of Competition Law and Economics*, 18(3):605–642.

Harrington, J. E. (2022b). Third-party algorithms and tacit collusion. *European Economic Review*, 142:103993.

Hart, S. and Mas-Colell, A. (2000). A simple adaptive procedure leading to correlated equilibrium. *Econometrica*, 68(5):1127–1150.

Hartline, J. D., Long, S., and Zhang, C. (2024). Regulation of algorithmic collusion. *arXiv preprint arXiv:2401.15794*.

Hettich, M. (2021). Algorithmic collusion: Insights from deep learning. *Available at SSRN 3785966*.

Holtz, D., Carterette, B., Chandar, P., Nazari, Z., Cramer, H., and Aral, S. (2020). The engagement-diversity connection: Evidence from a field experiment on spotify. In *Proceedings of the 21st ACM Conference on Economics and Computation*, pages 75–76.

Jannach, D. and Jugovac, M. (2019). Measuring the business value of recommender systems. *ACM Transactions on Management Information Systems*, 10(4):1–23.

Johnson, J. P., Rhodes, A., and Wildenbeest, M. (2023). Platform design when sellers use pricing algorithms. *Econometrica*, 91(5):1841–1879.

Kastius, A. and Schlosser, R. (2022). Dynamic pricing under competition using reinforcement learning. *Journal of Revenue and Pricing Management*, 21(1):50–63.

Keppo, J., Li, Y., Tsoukalas, G., and Yuan, N. (2025). AI Pricing, Agent Heterogeneity, and Collusion. Working paper, available at SSRN.

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

Koirala, P. and Laine, F. (2024). Algorithmic collusion in a two-sided market: A rideshare example. *arXiv preprint arXiv:2405.02835*.

Lamba, R. and Zhuk, S. (2022). Pricing with algorithms. *arXiv preprint arXiv:2205.04661*.

Lambin, X. (2023). Less than meets the eye: simultaneous experiments as a source of algorithmic seeming collusion. *Available at SSRN 4498926*.

Levine, D. K. (2024). Effciently breaking the folk theorem by reliably communicating long term commitments. *wp*.

Martin, S. and Rasch, A. (2024). Demand forecasting, signal precision, and collusion with hidden actions. *International Journal of Industrial Organization*, 92:103036.

Maskin, E. and Tirole, J. (1988). A theory of dynamic oligopoly, ii: Price competition, kinked demand curves, and edgeworth cycles. *Econometrica: Journal of the Econometric Society*, pages 571–599.

Meylahn, J. M. (2023). Does an intermediate price facilitate algorithmic collusion? *Available at SSRN 4594415*.

Miklós-Thal, J. and Tucker, C. (2019). Collusion by algorithm: Does better demand prediction facilitate coordination between sellers? *Management Science*, 65(4):1552–1561.

Milgrom, P. and Robert, J. (1990). Rationalizability, learning, and equilibrium in games with strategic complementarities. *Econometrica*, 39(1):1255–1277.

Musolff, L. (2022). Algorithmic pricing and dynamic market adjustment: Evidence from amazon. *Working Paper*.

Nikandrova, A. and Parekh, A. (2025). Algorithmic exclusion by large language models. Working paper, City St George's, University of London.

OECD (2021). Algorithms and collusion – competition policy in the digital age. In *OECD Competition Committee Reports*. Organisation for Economic Co-operation and Development. Accessed: 2025-07-11.

Orbach, B. (2024). Do revenue management platforms like realpage facilitate illegal algorithmic collusion? https://www.promarket.org/2024/04/18/do-revenue-management-platforms-like-realpage-facilitate-illegal-algorithmic-collusio Accessed: 2025-07-11.

O'Connor, J. and Wilson, N. E. (2021). Reduced demand uncertainty and the sustainability of collusion: How ai could affect competition. *Information Economics and Policy*, 54:100882.

Padilla, J., Perkins, J., and Piccolo, S. (2022). Self-preferencing in markets with vertically integrated gatekeeper platforms. *The Journal of Industrial Economics*, 70(2).

Peitz, M. and Sobolev, A. (2026). Inflated recommendations. *RAND Journal of Economics*. Forthcoming. CEPR Discussion Paper No. DP17260.

Popescu, D. (2015). Repricing algorithms in e-commerce.

Possnig, C. (2023). *Reinforcement learning and collusion*. Department of Economics, University of Waterloo.

Qiu, L., Huang, Y., Singh, P. V., and Srinivasan, K. (2025). Personalization, consumer search and algorithmic pricing. *Marketing Science*, 44(6):1278–1298.

Reimers, I. and Waldfogel, J. (2023). A framework for detection, measurement, and welfare analysis of platform bias. Working Paper 31766, National Bureau of Economic Research.

Ricci, F., Rokach, L., and Shapira, B., editors (2022). *Recommender Systems Handbook*. Springer, 3 edition.

Rocher, L., Tournier, A. J., and de Montjoye, Y.-A. (2023). Adversarial competition and collusion in algorithmic markets. *Nature Machine Intelligence*, 5(5):497–504.

Rubinstein, A. (1986). Finite automata play the repeated prisoner's dilemma. *Journal of economic theory*, 39(1):83–96.

Salcedo, B. (2015). Pricing algorithms and tacit collusion. *Manuscript, Pennsylvania State University*.

Schlechtinger, M., Kosack, D., Krause, F., and Paulheim, H. (2024). By fair means or foul:

Quantifying collusion in a market simulation with deep reinforcement learning. *arXiv preprint arXiv:2406.02650*.

Spann, M., Bertini, M., Koenigsberg, O., Zeithammer, R., Aparicio, D., Chen, Y., Fantini, F., Jin, G. Z., Morwitz, V. G., Popkowski Leszczyc, P., Vitorino, M. A., Williams, G. Y., and Yoo, H. (2025). Algorithmic pricing: Implications for marketing strategy and regulation. *International Journal of Research in Marketing*.

The Wall Street Journal (2025). Walmart is preparing to welcome its next customer: the ai shopping agent. Accessed: 2025-07-11.

Tufts Now (2025). Ai personal assistants could buy your groceries and book your plane tickets. Accessed: 2025-07-11.

U.S. Department of Justice (2024). Justice department sues realpage for algorithmic pricing scheme that harms millions of american renters. https://www.justice.gov/archives/opa/pr/justice-department-sues-realpage-algorithmic-pricing-scheme-harms-millions-american-r Accessed: 2025-07-11.

Wang, Q., Huang, Y., Singh, P. V., and Srinivasan, K. (2023). Algorithms, artificial intelligence and simple rule based pricing. *Available at SSRN 4144905*.

White House Council of Economic Advisers (2024). The cost of anticompetitive pricing algorithms in rental housing. https://bidenwhitehouse.archives.gov/cea/written-materials/2024/12/17/the-cost-of-anticompetitive-pricing-algorithms-in-rental-housing/. Accessed: 2025-07-11.

Wu, Z., Zheng, S., Liu, Q., Han, X., Kwon, B. I., Onizuka, M., Tang, S., Peng, R., and Xiao, C. (2024). Shall we talk: Exploring spontaneous collaborations of competing llm agents. *arXiv preprint arXiv:2402.12327*.

Xu, Z., Zhang, M., and Zhao, W. (2024). Algorithmic collusion and price discrimination: The over-usage of data. *arXiv preprint arXiv:2403.06150*.

Xu, Z., Zhang, M., and Zhao, W. (2025). Artificial intelligence, data and competition. *Working paper*.