NBER WORKING PAPER SERIES

ALGORITHMIC COERCION WITH FASTER PRICING

Zach Y. Brown
Alexander MacKay

Working Paper 34070
http://www.nber.org/papers/w34070

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2025, Revised January 2026

We thank Jim Dana, William Fuchs, Maxim Engers, Joe Harrington, Scott Kominers, Matt Leisten, Nathan Miller, and Andrew Sweeting for helpful comments. We thank participants at the Northwestern Antitrust Conference, Utah Winter Business Economics Conference, International Industrial Organization Conference, Barcelona Summer Forum Digital Economics Workshop, ZEW Conference on Economics of Information and Communication Technologies, and the Federal Trade Commission. We are grateful for the research assistance of Harry Kleyer. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Algorithmic Coercion with Faster Pricing
Zach Y. Brown and Alexander MacKay
NBER Working Paper No. 34070
July 2025, Revised January 2026
JEL No. D43, L13, L40, L81, L86

ABSTRACT

We study price competition when one firm uses a pricing algorithm that can react quickly to a rival's price. We characterize a coercive equilibrium in which the algorithm encodes a persistent rule that maximizes discounted profits subject to the rival's incentive compatibility constraint. The combination of faster pricing and multi-period commitment allows the algorithmic firm to unilaterally induce supracompetitive prices even when the rival is myopic and cannot sustain collusion. The algorithmic firm can earn more than its full collusion profits, and consumer surplus can fall below the full collusion benchmark. When the rival uses a learning rule to set prices, we show via simulations that outcomes rapidly converge to the coercive equilibrium. Finally, we demonstrate the implications of our framework for platform design.

Zach Y. Brown
University of Michigan
Department of Economics
and NBER
zachb@umich.edu

Alexander MacKay
University of Virginia
Department of Economics
mackay@virginia.edu

1 Introduction

In standard models of collusion, prices above competitive levels are sustained by the threat of punishment in future periods. Because all firms must be forward-looking and choose dynamic strategies that condition on the history of play, collusive outcomes are fragile and can be undermined by a single firm that maximizes current profits. For example, in a simultaneous price-setting game with differentiated products, a single firm facing short-sighted rivals cannot profitably price above the competitive (Bertrand–Nash) level in equilibrium. The framework, in which all firms coordinate on dynamic punishment strategies, provides the lens through which researchers and policymakers typically think about supracompetitive prices.[1]

However, with pricing algorithms, other forms of supracompetitive pricing are possible and perhaps widespread. In many markets, including retail gasoline, hotels, food delivery, and online retail, firms delegate pricing to software that monitors competitors and automatically adjusts prices several times per day or even per hour (e.g., Byrne et al., 2025; Aparicio et al., 2024; Brown and MacKay, 2023). Faster pricing and automated responses to rival prices are features explicitly advertised by pricing algorithm providers[2] and are defining characteristics of proprietary algorithms, including those employed by Amazon, the largest e-commerce retailer.[3] The increasing use of algorithms and the belief that the technology can enable new pricing strategies have raised concerns among policymakers worldwide.[4]

In this paper, we investigate the extent to which a single firm with a pricing algorithm can enforce supracompetitive prices, even when rivals maximize static profits and cannot collude. Our model has one firm (the *algorithmic firm*) commit to a pricing rule that specifies how it will respond to its rival's price. The algorithm enables more frequent price adjustments, providing an avenue for within-period punishment. The rival does not use an algorithm; instead, it simply chooses prices to maximize current-period profits. This is a conservative assumption that rules out voluntary coordination and tacit collusion. By showing that the algorithmic firm can coerce even a rival with no concern for future punishment, we establish the robustness of the mechanism. In this context, we characterize the *coercive equilibrium* in which the algorithmic firm can single-handedly implement prices substantially higher than competitive levels. In the

[1] Several additional factors are viewed as important for facilitating collusion, including similarity in size and costs, predictability of demand, observability of all rivals' prices, and the possibility of frequent direct communication. See, e.g., Tirole (1988) or Porter (2005).

[2] For instance, one provider of pricing tools for online retailers states that their pricing algorithm "helps you react instantly to a competitor's price move" (Retail Cloud, 2025). Another offers a premium service that "reacts to changes your competitors make in 90 seconds" (Repricer.com, 2023). Providers of pricing tools for offline markets advertise similar capabilities, such as software for gasoline stations that "automates your process for tracking competitive fuel prices" and updates in near real time (Taiga Data, 2025).

[3] Amazon's first-party algorithm can detect online price changes for thousands of popular products within hours and automatically respond (Federal Trade Commission et al., 2023).

[4] In the U.S., the Federal Trade Commission and the Department of Justice have alleged that pricing algorithms lead to higher prices in antitrust complaints. In addition, the European Commission and UK Competition and Markets Authority have launched investigations into algorithmic pricing practices.

limiting case, coercion can be worse for consumers than if firms explicitly coordinated on high prices.

We formalize the algorithmic firm's pricing technology using two primitives. First, within each period, the algorithm can respond to the rival's price with timing governed by a *speed advantage* parameter, which captures either more frequent price updates or more timely monitoring of rivals' prices. Second, the firm may not be able to revise its algorithm in every period. This is governed by a *commitment* parameter that captures the probability that the current algorithm persists. The product of this parameter and the firm's discount factor yields the firm's effective commitment; when it is high, the firm internalizes the fact that it will use the same algorithm for many periods.

Given this technology, we define a coercive equilibrium as follows. The algorithmic firm chooses an algorithm—equivalently, a target price pair and an associated punishment rule—to maximize its discounted profits, taking into account the rival's best response in each period. The rival, in turn, best responds to the algorithm in every period. Coercion arises because the algorithm ensures that deviations by the rival are met by a within-period response that sharply reduces the rival's profit. Rather than relying on the threat of future punishment, as in standard dynamic strategies of collusion, the algorithm reshapes the rival firm's within-period profits. In our baseline analysis, we assume that off the target path the algorithmic firm plays its one-shot best response to the rival's price. Under regularity conditions, the coercive equilibrium with this punishment is unique.

The model nests two familiar benchmarks. If there is no speed advantage but full commitment across periods, the algorithmic firm behaves as a leader in a sequential-move (Stackelberg pricing) game. If the algorithm can react instantaneously but cannot commit across periods, the algorithmic firm behaves instead as a follower.[5] Our interest lies in the interaction of these two features. We show that when the algorithm has both a speed advantage and multi-period commitment, the algorithmic firm can unilaterally implement substantially higher prices than under either benchmark.

Our results characterize how equilibrium prices and profits vary with the speed advantage and the degree of commitment. Even when its speed advantage and commitment are modest, the algorithmic firm can force its rival to choose a price strictly above the Bertrand price and can earn more than it would as a sequential leader. When the algorithm can respond almost instantaneously and commitment is strong, the algorithmic firm can induce the rival to charge a price above the fully collusive price, while itself setting a slightly lower price. In this *maximal coercion* case, the algorithmic firm earns more than its full collusion profits. Average prices can exceed collusive prices, leading to lower welfare and consumer surplus than that obtained with full collusion.

[5] In earlier work, we considered this latter case and the general possibility of a speed advantage to increase prices in Markov perfect equilibrium (Brown and MacKay, 2023). Our previous work did not address multi-period commitment.

The analysis highlights the complementarities between speed and commitment. A speed advantage alone allows the fast firm to behave as a follower, softening competition and raising prices. Multi-period commitment to an algorithm alone raises prices by allowing the committed firm to behave as a leader. When both features are present, the algorithmic firm can both discipline its rival through rapid punishment and discipline itself through commitment,[6] thereby transforming the feasible payoff set. In this sense, pricing algorithms that combine faster reaction and persistence of rules can be more powerful than either feature in isolation.

We show that the coercive mechanism is robust to alternative assumptions. First, allowing the rival to be forward-looking does not eliminate coercion: there exists an equilibrium in which the rival continues to play a static best response each period, so the coercive outcome remains an equilibrium even though additional equilibria may arise. Second, we compare our baseline best-response punishment to a price-matching rule under which the algorithm instantly matches any cut by the rival. With sufficiently fast adjustments, price matching can sustain fully collusive prices, though our baseline punishment assumption may be preferred by the algorithmic firm because it shifts surplus in its favor. Third, we show that the same equilibrium outcomes arise under alternative timing assumptions. Finally, we consider an oligopoly with one algorithmic firm and multiple slower rivals. Prices under maximal coercion fall as the number of non-algorithmic firms increases but remain well above Bertrand levels even when there are many rivals.

Our assumption that the rival is short-sighted can capture organizational constraints, legacy pricing systems, or lack of sophistication, as imposing dynamic strategies involves additional complexity (Fudenberg et al., 1990; Gale and Sabourian, 2005). A key question for practical relevance is whether coercion can arise when the rival does not understand that it faces an algorithm implementing a sophisticated punishment strategy. To address this, we introduce an extension in which the rival without an algorithm observes only its own prices and profits and updates its price using a simple gradient learning rule. In a simultaneous pricing game, such learning converges to the Bertrand–Nash equilibrium. We show that when the algorithmic firm commits to a linear pricing rule in the rival's price, the same learning process converges instead to a coercive equilibrium with supracompetitive prices. Intuitively, by reacting quickly to the rival's price changes, the algorithm can make the rival's perceived residual demand appear more inelastic, thereby guiding the rival toward higher prices.[7] We show that convergence to coercive outcomes is rapid on average and robust to demand shocks.

Our framework can also be applied to questions of platform design. Large online platforms can choose how often sellers may update prices and which pricing tools they may use. We model this by endogenizing the parameters governing speed and commitment by letting a platform

[6] Algorithm providers also emphasize this benefit of commitment, implying that this can safeguard against "rash" or overly short-sighted pricing decisions (FeedbackWhiz, 2025; Informed Software, 2025).

[7] This mechanism has a flavor of strategic manipulation of beliefs similar to signal-jamming models in dynamic games (Fudenberg and Tirole, 1986).

choose these features for the sellers on its marketplace. Platforms that place a high enough weight on seller profits have an incentive to equip some sellers with algorithms, even though doing so harms consumers. Enabling algorithms softens competition among sellers and raises profits even if the platform does not coordinate their pricing directly. When the platform is vertically integrated and sells its own products, it has a particularly strong incentive to grant itself the most powerful algorithmic technology. By contrast, a platform that puts equal weight on consumer surplus and the profits of competing sellers prefers to forbid algorithms that confer either a speed advantage or long-run commitment.

Our analysis contributes to several strands of the literature on collusion and algorithmic pricing. A growing theoretical literature studies how commitment by pricing algorithms can affect competition (Salcedo, 2015; Leisten, 2024; Levine, 2024). Lamba and Zhuk (2025) consider alternating-move settings in which commitment to algorithmic rules can raise prices above competitive levels. Musolff (2024) documents that third-party sellers on Amazon appear to commit to pricing rules that undercut rivals and then reset to higher prices. Despite the fact that the ability to quickly respond to competitors' price changes is a defining feature of algorithms, there is little theoretical work examining the role of algorithm speed. One exception is Brown and MacKay (2023), who analyze Markov-perfect outcomes when firms differ in how frequently they can adjust prices. In contrast, this paper is the first to explicitly analyze the implications when a speed advantage is combined with multi-period commitment, which substantially expands the set of feasible outcomes.

A parallel line of work investigates whether autonomous pricing algorithms can facilitate collusion in symmetric environments. Calvano et al. (2020), Asker et al. (2024), and Banchio and Mantegazza (2022) study when reinforcement-learning algorithms learn to sustain high prices in standard repeated games.[8] Hansen et al. (2021) study symmetric firms using misspecified learning algorithms. Miklós-Thal and Tucker (2019) and O'Connor and Wilson (2021) analyze how better demand forecasts affect the sustainability of collusion. Our contribution is complementary. Rather than asking whether symmetric algorithms can jointly learn to collude, we examine how a single firm can coerce its rivals through an algorithm that provides a speed advantage and commitment.

Empirical evidence provides further motivation. A literature documents widespread use of algorithmic pricing on online platforms, where sellers automate their responses to rivals' prices (Chen et al., 2016; Wieting and Sapi, 2021; Musolff, 2024). Assad et al. (2024) show that adoption of pricing algorithms in the German retail gasoline market coincides with more frequent price changes and higher prices. Perhaps most directly, Byrne et al. (2025) exploit a legal settlement that required one gasoline retailer to slow its response to rivals' prices. While their setting involves additional market complexities, the broad findings—that asymmetries

[8]A smaller literature has focused on alternating-move games as in Maskin and Tirole (1988). In particular, Klein (2021) examines collusion with machine learning algorithms in a sequential game.

in speed led to higher prices—are consistent with Brown and MacKay (2023) and with the implications of this paper.

Finally, we relate to recent work on platforms and third-party pricing algorithms. Johnson et al. (2023) examine how platforms can use demand-steering policies to increase competition when sellers use algorithms that collude. Harrington (2022), Calder-Wang and Kim (2023), Harrington (2025), and Sugaya et al. (2025) address the issues raised when multiple firms in a market adopt third-party pricing tools or engage in information sharing. We instead focus on how the design of a platform's proprietary algorithm (i.e., speed and commitment to a pricing rule) shapes competitive outcomes and platform incentives.

We argue that these seemingly benign features—faster reactions and persistent rules—can fundamentally change competitive outcomes. Unlike standard repeated-game models of collusion, which require all firms to be forward-looking and to understand the dynamic structure of the game, our coercive equilibrium persists even when rivals maximize static profits or follow simple learning rules.[9] High-speed pricing algorithms thus create a new channel through which firms can raise prices without explicit coordination.

The paper proceeds as follows. We introduce the model and equilibrium in Section 2. In Section 3, we discuss benchmark cases. We examine prices and welfare in Section 4. In Section 5, we introduce the idea of learning, and we show how the algorithmic firm can obtain coercive equilibria even with simple linear rules. Section 6 examines the incentives for platform design. Section 7 concludes.

2 The Model

We study a duopoly in which one firm uses a high-speed pricing algorithm and the other sets prices directly. Time is continuous but decisions are organized into discrete periods. The key asymmetry is that the algorithmic firm can commit (imperfectly) to a pricing rule across periods that reacts more quickly within a period. This section lays out the environment, the technology of the algorithmic firm, the behavior of the rival, and the equilibrium notion used in the rest of the paper.

2.1 Baseline Duopoly Environment

Two firms, a and b, each sell a single product with prices p_a and p_b. Firm a is the *algorithmic* firm while firm b sets prices directly without an algorithm. Time is continuous and is given by

[9] It is well known that in simultaneous-move games, high-frequency pricing implies a larger per-period discount factor, making collusion easier to sustain when firms have perfect monitoring (e.g., Abreu et al., 1991). Yet, all firms must be forward-looking to sustain collusion in these models. In contrast, we make the conservative assumption that the rival simply maximizes current-period profits.

$s \in [0, \infty)$, and it is indexed by discrete periods $t \in \{0, 1, 2, ...\}$. Period t corresponds to the continuous time interval $s \in [t, t + 1)$.

Demand arrives in continuous time within each period. Let x_t be a (possibly vector-valued) state that summarizes demand and cost conditions in period t. When prices are (p_a, p_b) and the state is x_t, the instantaneous profit flow of firm $i \in \{a, b\}$ is

$$\pi_i(p_i, p_{-i}, x_t), \tag{1}$$

which we assume is twice continuously differentiable and strictly concave in p_i for each (p_{-i}, x_t). To economize on notation, we abstract away from x_t and write $\pi_i(p_i, p_{-i})$. Throughout, x_t can be interpreted as an exogenous payoff shifter (e.g., demand or cost) observed at the start of each period. Since the evolution of x_t is exogenous, the equilibrium conditions and comparative statics below can be read as holding conditional on the current state.

We impose standard assumptions for differentiated-products price competition. Specifically, products are substitutes and prices are strategic complements:

$$\frac{\partial \pi_i(p_i, p_{-i})}{\partial p_{-i}} > 0 \quad \text{and} \quad \frac{\partial^2 \pi_i(p_i, p_{-i})}{\partial p_i \, \partial p_{-i}} > 0 \tag{2}$$

for $i \in \{a, b\}$. We further assume prices lie in compact intervals. Since π_i is strictly concave in own price, firm i's static best-response function $R_i(\cdot)$ exists, is single-valued, and is continuous.

The period length is chosen so that firm b can update its price at the beginning of each period. Firm a can also revise its price at that time, but, through its pricing algorithm, it can additionally adjust its price once more within the period, as described next.

2.2 Algorithmic Firm and Pricing Technology

Firm a is forward-looking and discounts future per-period profits with discrete discount factor $\beta \in [0, 1)$.[10] The pricing technology of firm a is summarized by an algorithm and characterized by two parameters:

- a *speed advantage* $\alpha \in [0, 1]$, which governs the fraction of the period during which the algorithm can exploit an updated price after observing p_b; and

- a *commitment parameter* $\gamma \in [0, 1]$, which governs how long a given algorithm remains in place.

[10]For simplicity, we assume there is no within-period discounting, consistent with the idea that the period is very short in many real-world settings with pricing algorithms. However, it is straightforward to redefine α, the speed advantage parameter defined below, as a parameter that accounts for within-period discounting. Consider any α and any instantaneous within-period discount rate ν. Then there exists an objective (non-discounted) speed advantage $\tilde{\alpha}$ such that $\frac{\int_0^{1-\tilde{\alpha}} e^{-\nu s} ds}{\int_0^1 e^{-\nu s} ds} = \frac{1-e^{-\nu(1-\tilde{\alpha})}}{1-e^{-\nu}} = 1 - \alpha$. For a given ν, the mapping of α to $\tilde{\alpha}$ is one-to-one.

At each reprogramming date τ, firm a's algorithm and firm b's price are chosen simultaneously. The algorithm $\mathcal{A}_\tau = (\rho_\tau, \sigma_\tau)$ consists of an initialization rule ρ_τ that sets a's price at the start of each period, and an update rule σ_τ that maps firm b's period-t price into an updated price for firm a. Formally, within any period t in which $\mathcal{A}_\tau$ is in force,

$$p_{as} = \mathcal{A}_\tau(p_{bt}, x_t, s) \equiv \begin{cases} \rho_\tau(x_t) & \text{if } s - t \in [0, 1-\alpha), \\ \sigma_\tau(p_{bt}, x_t) & \text{if } s - t \in [1-\alpha, 1), \end{cases} \tag{3}$$

for a candidate p_{bt}, where s is the index of continuous time. When $\alpha = 0$ the update window vanishes, and the model collapses to the simultaneous choice of ρ_τ and p_{bt}. When $\alpha = 1$ the update rule applies immediately after firm b chooses its price. We express the algorithm as a function of p_{bt} and not p_{bs} because firm b sets price only once per period.

There are two natural interpretations of this timing. One is that firms differ in how frequently they can change prices, even if they observe each other's price changes immediately. In this case, only the first opportunity for firm a to adjust after firm b moves is relevant, and the speed parameter α is a function of the ratio of firm b's price-change frequency to firm a's. Alternatively, firms may be able to update prices continuously but observe rivals at different frequencies. Under this interpretation, α is the fraction of the period during which firm a can adjust its price after observing firm b's price, before firm b can adjust again.

We assume that firm a modifies its algorithm infrequently. Firm a chooses an initial algorithm $\mathcal{A}$ at $t = 0$. Subsequently, at the beginning of each period, with probability γ the previous algorithm remains in force and cannot be changed; with complementary probability $1 - \gamma$ the firm can re-optimize and select a new algorithm. Thus γ captures the degree of (imperfect) commitment to a pricing rule across periods.[11]

Given these primitives, the algorithmic firm's problem can be expressed in terms of two value functions. Let $V_0(t)$ denote the value at the start of period t when firm a is able to revise its algorithm (for any p_{bt}). Let $V_1(t, \mathcal{A})$ denote the value when firm a is committed to algorithm $\mathcal{A}$ at the start of period t and anticipates that firm b best responds to that algorithm in future periods. The dynamic program can then be written as

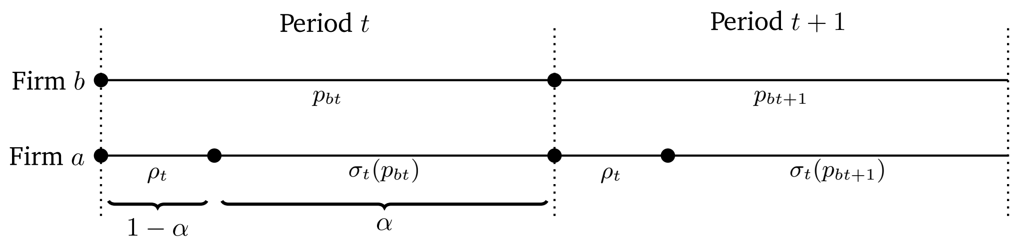
$$V_0(t) = \max_{\mathcal{A}} \Big[(1-\alpha)\, \pi_a(\rho, p_{bt}) + \alpha\, \pi_a(\sigma(p_{bt}), p_{bt}) + \beta\gamma\, V_1(t+1, \mathcal{A}) + \beta(1-\gamma)\, V_0(t+1)\Big] \tag{4}$$

$$V_1(t, \mathcal{A}) = (1-\alpha)\, \pi_a(\rho, p^*_{bt}) + \alpha\, \pi_a(\sigma(p^*_{bt}), p^*_{bt}) + \beta\gamma\, V_1(t+1, \mathcal{A}) + \beta(1-\gamma)\, V_0(t+1), \tag{5}$$

starting from $t = 0$ and for each integer t thereafter. Here p^*_{bt} is the optimal reaction of firm b

[11] We assume an exogenous commitment parameter to capture technological or institutional constraints on reprogramming frequency. We then endogenize this choice in Section 6. This structure provides a foundation for analyzing other frictions that give rise to commitment to a pricing rule, including Poisson arrival processes or fixed costs per update.

Figure 1: Timing with Differences in Pricing Speed



Notes: The figure illustrates the within-period timing of pricing when firm a, the algorithmic firm, sets algorithm $\mathcal{A}_t^{\alpha,\gamma}$ at the start of period t and remains committed to the algorithm in period $t+1$. The example depicts $\alpha = 3/4$ and the black dots are opportunities to adjust prices.

to algorithm $\mathcal{A}$, defined in the next subsection.

It is convenient to collect the present value of payoffs into the auxiliary function

$$\tilde{V}(t) \equiv \frac{1-\beta}{1-\beta\gamma} V_0(t). \tag{6}$$

With time-invariant profit functions the problem is stationary, so $V_0(t)$ and $V_1(t, \mathcal{A})$ do not depend on t. Combining (4) and (5), we can write firm a's optimization as

$$\begin{aligned} \tilde{V}(t) = \max_{\mathcal{A}} & \Big[(1-\alpha)\,\pi_a(\rho, p_{bt}) + \alpha\,\pi_a(\sigma(p_{bt}), p_{bt})\Big] \\ & + \sum_{r=t+1}^{\infty} (\beta\gamma)^{r-t} \Big[(1-\alpha)\,\pi_a(\rho, p_{bt}^*) + \alpha\,\pi_a(\sigma(p_{bt}^*), p_{bt}^*)\Big]. \end{aligned} \tag{7}$$

Figure 1 represents the within-period timing graphically and highlights how the algorithmic technology changes the intra-period timing relative to the benchmark with simultaneous price setting. When $\alpha = 0$ the lower line has no interior dots and the model collapses to simultaneous pricing. When $\alpha > 0$, firm a can undercut, match, or otherwise respond to firm b's price during the final fraction α of the period.

To make the speed advantage and commitment parameters concrete, consider the following example. Firm b sets its price once per week. Firm a has software that allows for daily price changes, and it updates the algorithm once per quarter on average. Then $\alpha = 6/7 \approx 0.86$ and $\gamma \approx 0.92$. An annual discount factor of 0.9 corresponds to a weekly discount factor $\beta = 0.998$ and yields $\beta\gamma \approx 0.92$. If firm a's software updates prices hourly instead, then $\alpha \approx 0.99$.

2.3 Non-Algorithmic (Slow) Rival

Firm b does not use an algorithm and is instead modeled as a static optimizer that chooses its price to maximize current-period profits. In equilibrium, firm b will best-respond to $\mathcal{A}_\tau$ in the current period but ignore any implications for future periods. This assumption greatly simplifies the dynamic structure and, as in the literature on long-run versus short-run players, gives a conservative benchmark for the scope of supracompetitive outcomes.[12]

Assumption A1 (Rival Maximizes Static Profits)**.** *In each period $t \in \{0, 1, 2, \ldots\}$, firm b solves*

$$\max_{p_b} (1-\alpha)\, \pi_b(p_b, \rho_\tau) + \alpha\, \pi_b\big(p_b, \sigma_\tau(p_b)\big), \tag{8}$$

where ρ_τ and σ_τ are the initialization and update rules embedded in the algorithm $\mathcal{A}_\tau$ chosen by firm a at the most recent reprogramming date τ.

The period length is defined so that firm b updates its price once per period. The objective in (8) therefore aggregates the profits earned in the portion of the period before the algorithm updates (weight $1-\alpha$) and in the portion after the update (weight α).

Assumption A1 provides a lower bound on the algorithm's coercive capability. In standard repeated games, the sustainability of supracompetitive prices relies on rivals valuing future profits enough to fear punishment. A myopic rival, by contrast, perceives no "continuation value" and, consequently, the algorithmic firm cannot rely on the rival's patience to sustain high prices; it must instead rely entirely on its technological advantage—speed and commitment—to make deviations unprofitable *within the period*. If the algorithm can coerce a myopic rival, it satisfies a stricter incentive constraint than that required to discipline a forward-looking one.

The assumption that the rival maximizes static profits is equivalent to assuming that a forward-looking rival plays a Markovian strategy. In a dynamic setting, if firm b does not condition its strategy on the history of play (i.e., it does not employ trigger strategies), its optimal policy is to treat the current algorithm $\mathcal{A}_\tau$ as a fixed constraint. Since firm b's current price choice p_{bt} does not influence the probability γ that the algorithm persists, firm b cannot strategically manipulate future market conditions. Therefore, maximizing current-period profits is the optimal dynamic strategy, and the coercive outcome is the unique Markov perfect equilibrium.

2.4 Coercive Equilibrium

An equilibrium consists of an algorithm for firm a and a pricing strategy for firm b such that each is optimal given the other. Formally, an equilibrium is a pair

$$\big(\{\mathcal{A}^*_\tau\}_{\tau \geq 0}, \{p^*_{bt}\}_{t \geq 0}\big) \tag{9}$$

where $\mathcal{A}^*_\tau$ denotes the algorithm used at reprogramming dates τ and $\{p^*_{bt}\}$ is the sequence of prices chosen by firm b, such that:

[12]This mirrors the classic literature on reputation with one long-run player facing a sequence of short-run opponents, e.g., Milgrom and Roberts (1982), Kreps and Wilson (1982), Fudenberg and Levine (1989), and Fudenberg et al. (1990).

1. Given $\{p^*_{bt}\}$, each $\mathcal{A}^*_\tau$ solves firm a's dynamic program (4)–(5) at $t = \tau$.

2. Given $\{\mathcal{A}^*_\tau\}$, each p^*_{bt} solves firm b's static problem (8) (Assumption A1).

With x_t constant, strategies are stationary, such that $\{\mathcal{A}^*_\tau\} = \mathcal{A}^*\ \forall \tau$.

For our baseline analysis we consider a class of update functions with best-response punishment. The algorithm does not implement a price change within the period if firm b chooses a particular *target price*. Otherwise, the algorithm updates prices according to a standard convention for punishment. Formally,

Assumption A2 (Best-Response Punishment)**.** *For a target price $p^\dagger_b$ chosen by firm b, $\sigma(p^\dagger_b) = \rho$. Otherwise, the algorithm uses the static best-response as punishment, i.e., $\sigma(\cdot) = R_a(\cdot)\ \forall p_b \neq p^\dagger_b$.*

While we use the term "punishment" to denote the off-path response $\sigma(\cdot)$, strictly speaking, a myopic rival perceives this not as a penalty for past actions, but as a kink in the mapping from its own price to current profits.

Under A1 and A2, it is convenient to re-express the algorithmic firm's problem in terms of a pair of *target prices*. Let $(p^\dagger_a, p^\dagger_b)$ denote the prices that firm a would like to be played on the equilibrium path in every period. We then interpret the algorithm as implementing these targets on the path and specifying a punishment off the path. In particular, given the current rival price p_{bt}, firm a chooses a target price pair and algorithm components (ρ, σ) that solve

$$\tilde{V}(t) = \max_{(p^\dagger_a, p^\dagger_b)} \left[(1-\alpha)\, \pi_a(\rho, p_{bt}) + \alpha\, \pi_a(\sigma(p_{bt}), p_{bt}) + \frac{\beta\gamma}{1-\beta\gamma}\, \pi_a(p^\dagger_a, p^\dagger_b) \right] \tag{10}$$

$$\text{s.t.} \quad (i) \quad \rho = p^\dagger_a, \tag{11}$$

$$(ii) \quad \sigma(p_b) = \begin{cases} p^\dagger_a & \text{if } p_b = p^\dagger_b, \\ R_a(p_b) & \text{if } p_b \neq p^\dagger_b, \end{cases} \tag{12}$$

$$(iii) \quad \pi_b(p^\dagger_b, p^\dagger_a) \ \geq\ (1-\alpha)\, \pi_b(\hat{p}_b, \rho) + \alpha\, \pi_b\big(\hat{p}_b, \sigma(\hat{p}_b)\big) \quad \forall \hat{p}_b, \tag{13}$$

which is obtained by plugging the target prices $(p^\dagger_a, p^\dagger_b)$ into (7). Constraint (i) requires the initialization rule to implement the target price for firm a on the equilibrium path. Constraint (ii) requires the update rule to deliver the same price on the path and to follow a specified punishment when firm b deviates. Constraint (iii) is the incentive compatibility condition for firm b, comparing the profits from following the target price to the profits from deviating to any $\hat{p}_b$ given the algorithm's response. We adopt a standard tie-breaking convention: when firm b is indifferent, it complies with the on-path target.

Under additional regularity conditions (which we relegate to the Appendix), we have existence and uniqueness of equilibria.

Proposition 1 (Equilibrium Existence and Uniqueness)**.** *Under Assumption A1, Assumption A2, and Appendix Assumption R1, a coercive equilibrium exists and is unique.*

All proofs are in Appendix A.1. Assumption A1 is a conservative assumption about conduct that limits the scope for supracompetitive pricing. Assumption A2 anchors off-path play and serves as an equilibrium selection device when considering all Markov perfect equilibria among the unconstrained set of punishment functions. Appendix Assumption R1 collects regularity conditions: (i) firm a's objective is single-peaked and (ii) the best-response slope for firm b is less than one (which yields a unique fixed point). These conditions are satisfied in standard differentiated-products demand environments and can be verified directly in our linear-demand illustrations. In addition, we impose (iii) a selection rule at the boundary where the maximal incentive-compatible target coincides with the current price. In Section 4.3, we explore robustness to these assumptions. We show that the coercive equilibrium exists (but need not be unique) if the rival is forward-looking, provide results for alternative punishment strategies, and consider alternative timing assumptions that preclude the boundary selection rule.

In the next section we use this equilibrium mapping to study how the algorithm's speed advantage α and commitment parameter γ interact to generate coercive outcomes and to compare these outcomes to the familiar simultaneous-move and sequential-move benchmarks.

3 Coercive Equilibrium: Benchmark Cases

To organize the analysis, it is useful to first examine a set of benchmark environments that isolate features of the algorithmic technology. We consider four cases. First, we study the environment in which the algorithm cannot deliver either a speed advantage or multi-period commitment. Second, we allow for commitment across periods but continue to assume that prices are updated simultaneously within each period. Third, we allow for a speed advantage within a period but rule out multi-period commitment. Finally, we combine the two dimensions and analyze the limiting case in which the algorithm is both arbitrarily fast and perfectly persistent; this is the maximal coercion benchmark that anchors our more general results in Section 4.

For comparison, we also refer to the outcome that maximizes joint profits, which we label the *collusive* outcome. This is the allocation that would arise if the two firms could operate as a perfect cartel or merge into a multi-product monopolist. Formally, the collusive prices solve

$$\max_{(p_a,p_b)} \; \pi_a(p_a,p_b)+\pi_b(p_b,p_a), \tag{14}$$

and we define each firm's *full collusion profits* as the profit it earns at these prices.

3.1 Simultaneous Pricing

We begin with the case in which the algorithmic technology delivers neither a speed advantage nor commitment. Setting $\alpha = 0$ and $\gamma = 0$, firm a's algorithm is effectively a static pricing rule, and both firms choose prices myopically each period. The objective functions reduce to

$$\text{Firm } a: \quad \max_{\mathcal{A}} \ \pi_a(\rho, p_{bt}), \tag{15}$$

$$\text{Firm } b: \quad \max_{p_b} \ \pi_b(p_b, \rho_\tau), \tag{16}$$

where ρ is the initial price chosen by firm a's algorithm in the period and ρ_τ denotes the corresponding initialization in period τ.

Because neither firm values future periods in its choice problem and there is no within-period speed advantage, this dynamic game collapses to the one-shot simultaneous price-setting game. The unique subgame-perfect equilibrium is therefore the standard Bertrand–Nash outcome. While a general dynamic pricing game may admit many equilibria, the presence of a rival that maximizes current-period profits sharply restricts the equilibrium set.

3.2 Commitment without Speed Advantage

We next allow the algorithmic firm to commit across periods but continue to assume that prices are updated simultaneously within each period.[13] This corresponds to the case $\alpha = 0$ and $\gamma > 0$. The algorithmic firm can commit to a rule that pins down its price in all future periods until the algorithm is reprogrammed. Under these assumptions, firm a's objective is

$$\max_{\mathcal{A}} \pi_a(\rho, p_{bt}) \ + \ \frac{\beta\gamma}{1-\beta\gamma}\, \pi_a\big(\rho, p^*_{bt}\big), \tag{17}$$

while firm b continues to solve

$$\max_{p_b} \ \pi_b(p_b, \rho_\tau). \tag{18}$$

Here $p^*_{bt} = R_b(\rho)$ is firm b's static best response to the initialization price ρ. The first term in (17) is the simultaneous price-setting incentive: firm a's short-run profit given p_{bt}. The second term captures the continuation value of committing to ρ, which induces firm b to adjust to p^*_{bt} in all future periods in which the algorithm remains in place.

It is convenient to interpret the product $\beta\gamma$ as an *effective commitment parameter*. When $\beta\gamma$ is close to zero, the continuation term is negligible and firm a behaves much like a static competitor. When $\gamma = 1$ and $\beta \to 1$, the continuation term dominates and firm a effectively behaves as a sequential leader, internalizing firm b's best response. We use $\beta\gamma = 1$ to denote this full commitment case (formally $\beta\gamma \to 1$). This corresponds to periods that are very short (so β is arbitrarily close to one) and pricing rules that are arbitrarily persistent (so γ is arbitrarily

[13] This is the case in the literature on learning algorithms and competition (e.g., Calvano et al., 2020; Asker et al., 2024; Johnson et al., 2023), in which firms commit to learning algorithms that set prices simultaneously. The analysis here can be roughly thought of as an extension of these models, where one firm can endogenously choose the optimal algorithm and the other firm uses a learning algorithm that recovers its true payoffs. We will not discuss learning here but instead describe the long-run payoffs.

close to one).

The resulting equilibrium relationship between prices can be summarized as follows.

Proposition 2 (Commitment Benchmark)**.** *Suppose the algorithm enables commitment across periods but no within-period speed advantage, so that $\beta\gamma > 0$ and $\alpha = 0$. Then the unique equilibrium price vector lies on firm b's static best-response function between the simultaneous price-setting (Bertrand) outcome and the sequential price-setting outcome in which firm a moves first.*

Proposition 2 shows that persistence alone can convert an otherwise simultaneous-move pricing problem into a sequential one. When $\alpha = 0$, the firms still choose prices at the same instant within each period, but firm a selects an algorithm whose initialization ρ is then locked in for all future periods in which the algorithm remains deployed. Because firm b behaves myopically, its continuation behavior collapses to the static best response $p^*_{bt} = R_b(\rho)$. Anticipating this, firm a chooses ρ to trade off the short-run incentive to best respond to the current p_{bt} against the continuation gains from committing and thereby inducing b to move to $R_b(\rho)$ going forward. As $\beta\gamma \to 1$, the continuation term dominates and firm a solves the Stackelberg leader problem $\max_\rho \pi_a(\rho, R_b(\rho))$, yielding the standard sequential-move outcome with a as the leader. When prices are strategic complements, this leader-like commitment motive pushes both prices above Bertrand–Nash. When the firms are symmetric, the leader sets the higher price because it internalizes how a higher commitment price shifts the follower's best response upward.

3.3 Speed Advantage without Commitment

We now turn off commitment and isolate the role of the within-period speed advantage. Set $\gamma = 0$ and let $\alpha > 0$, so that firm a can react more quickly to its rival's price but cannot commit across periods. In period t, firm a chooses an algorithm $\mathcal{A}$ that determines an initial price ρ for the first $(1-\alpha)$ fraction of the period, and an update rule $\sigma(\cdot)$ for the remaining fraction α. Its objective is

$$\max_{\mathcal{A}} \, (1-\alpha)\,\pi_a(\rho, p_{bt}) \; + \; \alpha\,\pi_a\big(\sigma(p_{bt}), p_{bt}\big), \tag{19}$$

while firm b continues to solve

$$\max_{p_b} \, (1-\alpha)\,\pi_b(p_b, \rho_\tau) + \alpha\,\pi_b\big(p_b, \sigma_\tau(p_b)\big), \tag{20}$$

taking the algorithm as given.

This problem is equivalent to the asymmetric commitment model analyzed by Brown and MacKay (2023). In that setting, it is weakly dominant for the faster firm to choose an update rule equal to its static best response, $\sigma(\cdot) = R_a(\cdot)$, which coincides with our Assumption A2. With this choice, firm a's problem in (19) trades off two forces: a simultaneous price-setting incentive over the first $(1-\alpha)$ portion of the period and a sequential follower incentive over the remaining portion, where firm a best responds to p_{bt}.

The resulting equilibrium lies along firm a's best-response function between the Bertrand outcome and the sequential equilibrium in which firm a is the follower. The parameter α governs the weight the slower firm places on the part of the period during which the algorithm has already reacted. As in the commitment-only case, when products are substitutes and prices are strategic complements, both firms' prices are higher than in the Bertrand equilibrium. Unlike the commitment-only benchmark, however, the algorithmic firm may price *below* its rival when the firms are symmetric, reflecting its follower role within the period.

3.4 Maximal Coercion

The benchmarks above isolate the role of commitment across periods and the within-period speed advantage. Our main interest lies in the interaction between the two. To highlight the forces at work, we consider a limiting case in which firm a has both maximal speed ($\alpha = 1$) and maximal commitment ($\beta\gamma = 1$). In this limit, firm a is arbitrarily patient, programs the pricing rule once and sets it to run forever, and has a pricing rule that responds instantaneously to any deviation by firm b.

Under maximal coercion, firm a chooses a target price pair $(p_a^\dagger, p_b^\dagger)$ and an algorithm that implements this pair whenever firm b follows the target and punishes deviations using the best-response from Assumption A2. Because commitment is perfect and the algorithm can respond instantaneously, the dynamic problem reduces to the static program

$$\begin{aligned} &\max_{(p_a^\dagger, p_b^\dagger)} \; \pi_a(p_a^\dagger, p_b^\dagger) \\ &\text{s.t.} \quad \pi_b(p_b^\dagger, p_a^\dagger) \;\geq\; \max_{p_b} \pi_b\big(p_b, R_a(p_b)\big). \end{aligned} \tag{21}$$

The constraint is firm b's incentive compatibility condition: given the punishment function $R_a(\cdot)$, the target price $p_b^\dagger$ must deliver at least as much profit as any deviation. With $R_a(\cdot)$ as punishment, the right-hand side of the constraint equals the profit that firm b would obtain as the leader in a sequential-move pricing game. Under maximal coercion, firm a therefore selects a point on firm b's isoprofit curve corresponding to its sequential-leader payoff and chooses the point on that curve which maximizes its own profit.

It is important to distinguish the mechanism of coercion here from standard repeated-game punishment. In our setting, the rival is myopic and ignores future consequences. Consequently, the "punishment" embedded in the algorithm does not function as a threat of future retaliation. Instead, the speed advantage α allows the algorithmic firm to reshape the rival's current-period profit function. By committing to a rule that reacts strictly within the period, the algorithm effectively constructs a synthetic residual demand curve for the rival. If the rival cuts its price below the target $p_b^\dagger$, the algorithm responds immediately, rendering the price cut unprofitable for that period. The rival, optimizing against this immediate response, finds that maintaining the high target price is a static best response.

In the next subsection we illustrate how far the maximal coercion benchmark can be from the Bertrand and sequential-move outcomes in a linear-demand setting.

3.5 Numerical Example with Linear Demand

We illustrate our results with examples generated from a symmetric linear demand system given by

$$D_i(p_i, p_{-i}) = 1 - \left(\frac{1}{4} + \frac{d}{2}\right) p_i + \frac{d}{2} p_{-i} \tag{22}$$

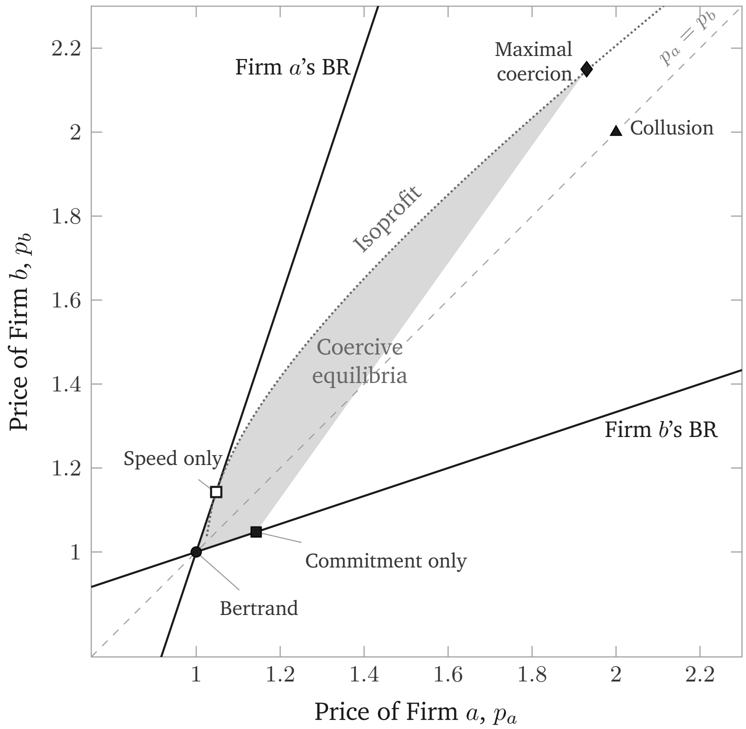
where $d > 0$ is an inverse measure of product differentiation. This demand system can be derived from a quasilinear quadratic utility function (Singh and Vives, 1984). Without loss of generality, we normalize marginal costs to zero. We parameterize demand such that total profits under joint-profit maximization are always 2 for any level of differentiation.

Figure 2 plots four benchmark price pairs for the linear-demand in (22) with $d = 1$ in (p_a, p_b)-space. The plotted values capture the four limiting cases with either no or full commitment ($\beta\gamma \in \{0, 1\}$) and either no speed advantage or an immediate reaction ($\alpha \in \{0, 1\}$). The best-response functions and the 45-degree line are shown for reference. With neither a speed advantage nor commitment, the unique outcome is the symmetric Bertrand–Nash equilibrium, $(p_a, p_b) = (1, 1)$. With only full commitment and no speed advantage, the equilibrium is equivalent to the sequential-leader benchmark, $(p_a, p_b) = (1.14, 1.05)$. A speed advantage and no commitment flips the firm's roles and produces the sequential-follower benchmark, $(p_a, p_b) = (1.05, 1.14)$. These latter two cases yield moderate price increases above Bertrand (5–14 percent).

When the algorithmic firm has maximum speed advantage and full commitment, $(\alpha, \beta\gamma) = (1, 1)$, it obtains the maximal-coercion outcome $(p_a, p_b) = (1.93, 2.15)$. This is the most profitable point consistent with the rival's deviation constraint (which is the isoprofit curve shown). These prices are far above the Bertrand benchmark, and the quantity-weighted average price is 2.01, above the collusive price level of 2. The algorithmic firm prices meaningfully below its rival, but both prices are now 93 and 115 percent higher than in the Bertrand benchmark—much higher than the sequential outcomes.

The algorithmic firm implements the maximal coercive outcome using an update rule that coincides with its static best response whenever firm b deviates from the target price. This is

Figure 2: Benchmark Equilibria



Notes: The figure illustrates benchmark cases under varying values of α and $\beta\gamma$, as described in the text. Bertrand corresponds to $(\alpha, \beta\gamma) = (0, 0)$, commitment only corresponds to $(\alpha, \beta\gamma) = (0, 1)$, speed only corresponds to $(\alpha, \beta\gamma) = (1, 0)$, and maximal coercion corresponds to $(\alpha, \beta\gamma) = (1, 1)$. The horizontal axis provides the price for firm a and the vertical axis provides the price for firm b. The gray shaded region illustrates the set of coercive equilibria over the full range of values for α and $\beta\gamma$. Best-response functions and the collusive outcome are shown for reference. The provided isoprofit curve indicates the points that provide firm b with the sequential-leader profit. Assumes $d = 1$ under the linear demand system in equation (22).

given by

$$\sigma(p_{bt}) = \begin{cases} 1.93 & \text{if } p_{bt} = 2.15, \\ (2 + p_{bt})/3 & \text{if } p_{bt} \neq 2.15, \end{cases} \tag{23}$$

which is simply firm a's best response to p_{bt} away from the target pair. If firm b sets $p_{bt} = 2.15$, its profit is $\pi_b(2.15, 1.93) = 0.76$. If it deviates and triggers the punishment, it faces the best-response function $(2+p_{bt})/3$ and solves $\max_{p_b} \pi_b\big(p_b, (2+p_b)/3\big)$, which yields the same profit as at the target.[14] Thus, the incentive compatibility constraint binds and the target pair $(1.93, 2.15)$ is a coercive equilibrium.

The implications for profits are substantial. In this example, firm a earns profits of 0.75 in the Bertrand equilibrium, 0.76 as the sequential leader, and 0.82 as the sequential follower. Under maximal coercion, it earns 1.21. Moving from Bertrand to maximal coercion therefore raises firm a's profits by about 61 percent. Moreover, the algorithmic firm earns strictly more

[14] Firm b's optimal deviation from the maximal coercion equilibrium is to choose a price of 1.14, which would imply that firm a's algorithm sets a price of 1.05. These are the sequential leader and follower prices.

than its full collusion profits: at the joint-profit-maximizing prices $(p_a, p_b) = (2, 2)$, each firm earns 1, whereas maximal coercion yields profits of 1.21 for firm a and only 0.76 for firm b. In this sense, coercion can be more profitable for the algorithmic firm than collusion, while leaving its rival strictly worse off.

The shaded set in Figure 2 depicts the full collection of coercive equilibrium price pairs that arise as $(\alpha, \beta\gamma)$ varies over $[0, 1]^2$. The next section characterizes this attainable set more systematically and studies how increases in speed and commitment expand the range of coercive outcomes.

4 Prices and Welfare in Equilibrium

4.1 Equilibrium Prices and Profits

We now study how the coercive equilibrium varies with the parameters of the algorithmic firm's pricing technology. From the previous section, we observe two striking properties of coercive equilibria:

Remark 1. *In the coercive equilibrium, the rival's price can exceed the price that maximizes joint profits.*

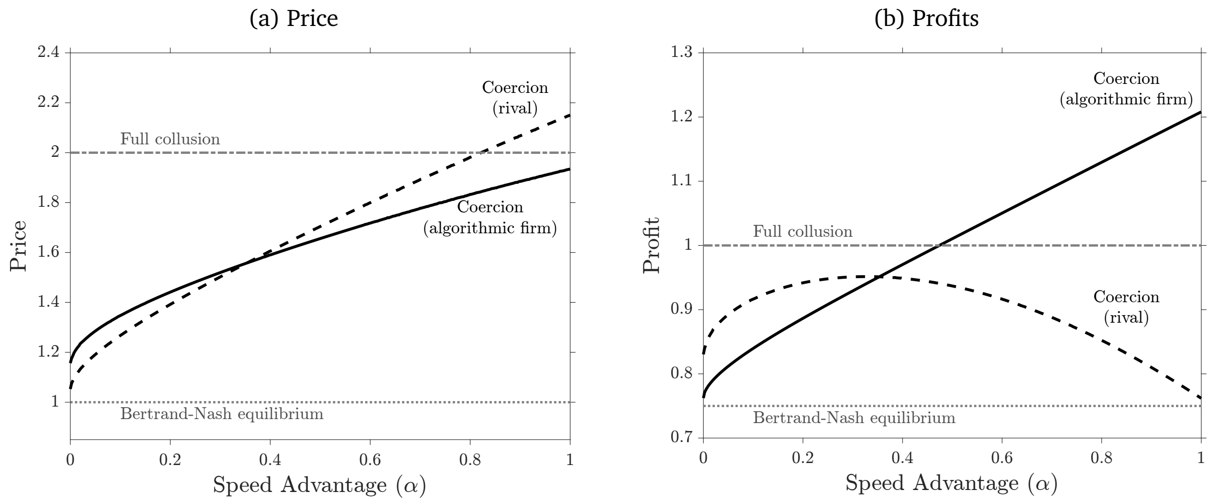
Remark 2. *In the coercive equilibrium, the algorithmic firm can earn more than its full collusion profits.*

Here, we show how these properties occur even with intermediate values of speed and commitment. Speed increases the strength of punishment, while commitment reduces the algorithmic firm's temptation to reduce prices for short-run gains. Together, these features expand the set of feasible target prices.

We first provide illustrations using the linear demand system in equation (22) with $d = 1$, so that the benchmark cases in Section 3.5 provide a convenient comparison. We then characterize general properties of the equilibrium with formal results.

Figure 3 reports equilibrium prices and profits when the algorithmic firm has full commitment ($\beta\gamma = 1$) and its speed advantage α varies from 0 to 1. The solid curve shows the outcome for the algorithmic firm, while the dashed curve shows the outcome for the slower rival. For reference, the figure also depicts the joint-profit maximizing ("collusive") outcome and the Bertrand benchmark with simultaneous pricing. Joint profits are maximized at $(p_a, p_b) = (2, 2)$, which yields profits of 1 for each product. As panel (a) of Figure 3 shows, when there is full commitment and α is sufficiently large (approximately $\alpha > 0.8$), the slower rival's equilibrium price lies strictly above the collusive level. Intuitively, the algorithmic firm exploits its speed advantage to induce the rival to set a high price and then attracts a disproportionate share of demand at a slightly lower price. For intermediate values of α (approximately $\alpha > 0.5$), the

Figure 3: Prices and Profits in Coercive Equilibrium, by Algorithmic Firm's Pricing Speed



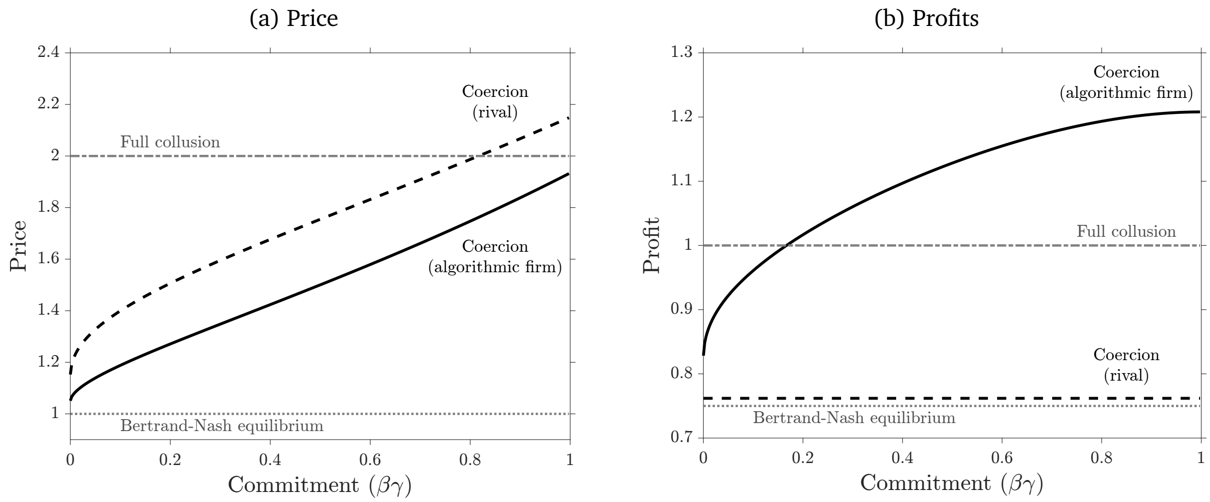
Notes: Panel (a) shows equilibrium prices and panel (b) shows equilibrium profits under Bertrand competition with simultaneous pricing, coercion with full commitment ($\beta\gamma = 1$), and joint profit maximization. The figure reports outcomes for different values of the speed advantage α of the algorithmic firm. Assumes $d = 1$ under the linear demand system in equation (22).

algorithmic firm already earns more than its full collusion profits, as seen in panel (b). In this case, joint profits are close to the collusive level, but the algorithmic firm obtains a larger share of those profits.

Figure 4 traces out the equilibrium when the algorithmic firm has the fastest possible reaction speed ($\alpha = 1$) and commitment $\beta\gamma$ varies. As commitment increases, both firms' prices and profits rise above the Bertrand benchmark. With weak commitment, the algorithmic firm is tempted to set a low initial price to respond to its rival's price, as there is a short-run gain and minimal long-run cost. This leads to lower prices in equilibrium. As $\beta\gamma$ increases, the algorithmic firm places more weight on future profit flows and can sustain a higher own price even when short-run incentives to undercut increase. In the limit of full commitment and maximum speed, the algorithmic firm is able to push the rival's price above the collusive level while keeping its own price slightly below, thereby capturing a large share of industry profits.

These figures illustrate how a speed advantage combined with commitment across periods effectively enlarges the feasible profit set relative to the standard Bertrand and sequential-move benchmarks. Profits are monotonically increasing in speed and commitment for the algorithmic firm, but have different implications for the rival firm. As shown in Figure 3, when the algorithmic firm's speed advantage is small, the rival earns greater profits than the algorithmic firm. Around $\alpha = 0.4$, the lines for firm a and firm b intersect. These values reflect the case when the incentives from commitment and the pricing advantage balance each other out, yielding symmetric prices and profits. By contrast, in Figure 4, profits for the rival never exceed the sequential-move level. When $\alpha = 1$, the algorithmic firm can punish more aggressively for any

Figure 4: Prices and Profits in Coercive Equilibrium, by Algorithmic Firm's Commitment



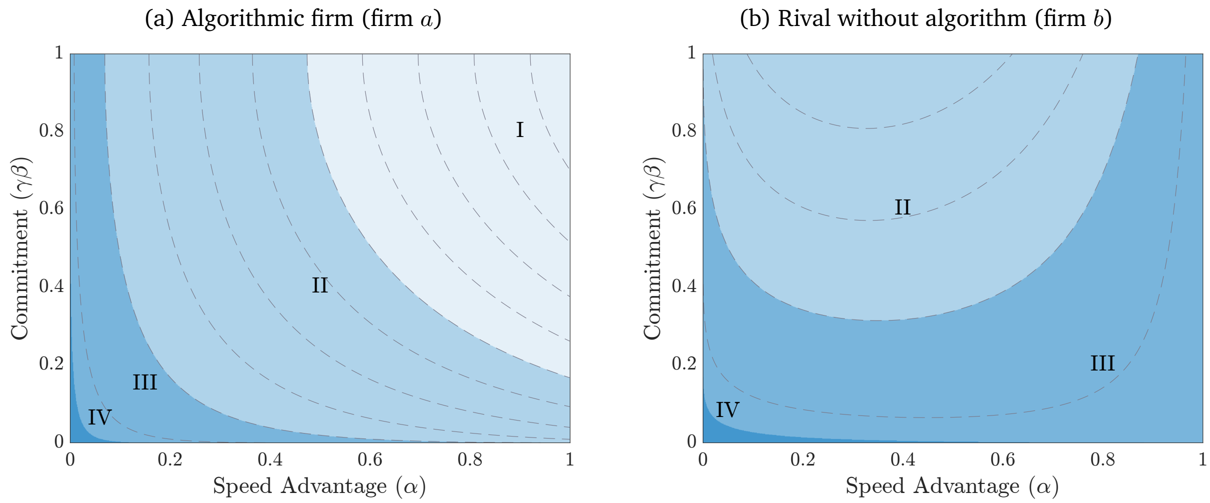
Notes: Panel (a) shows equilibrium prices and panel (b) shows equilibrium profits under Bertrand competition with simultaneous pricing, coercion with maximum speed ($\alpha = 1$), and joint profit maximization. The figure reports outcomes for different values of commitment $\beta\gamma$. Assumes $d = 1$ under the linear demand system in equation (22).

level of commitment.

For this example, equilibrium prices are supermodular in speed and commitment in the sense that the marginal effect of speed is larger when commitment is higher (and vice versa). With only commitment or only a speed advantage, the algorithmic firm is bounded in its ability to raise prices to the sequential prices and payoffs. Without commitment, the algorithmic firm has a temptation to accommodate the rival's price for short-term gain (behaving like a standard follower). Without speed, the firm can commit, but lacks the instrument to discipline the rival's within-period deviation. In the presence of both features—as illustrated by moving left to right in Figure 3 or Figure 4—an algorithmic firm can obtain substantially higher prices and profits. These complementarities are apparent at intermediate values of commitment and speed advantage, as illustrated in Appendix Figures A-3 and A-4. For example, the algorithmic firm can obtain profits greater than collusive profits if either $\beta\gamma = 0.5$ or $\alpha = 0.5$.

Figure 5 summarizes these comparative statics by plotting profits over the $(\alpha, \beta\gamma)$ space. Profit regions for the algorithmic firm are shown in panel (a) and those for the slower rival are shown in panel (b). Region I indicates profits above the collusive level, region II indicates profits between the collusive and sequential-follower benchmarks, region III indicates profits between the sequential-follower and sequential-leader benchmarks, and region IV indicates profits between the sequential-leader and Bertrand benchmarks. The algorithmic firm's profit exceeds the sequential-leader benchmark for almost all combinations of α and $\beta\gamma$, and there is a sizable region I in which its profit strictly exceeds the symmetric collusive level. The slower rival typically earns more than it would as a sequential leader, but never more than its collusive

Figure 5: Profit Regions by Pricing Speed and Commitment



Notes: The figure reports profit regions for firm a and firm b as a function of the algorithmic firm's speed advantage α and commitment $\beta\gamma$. Region I consists of parameter pairs for which profits are above a firm's full collusion profits; Region II corresponds to cases in which profit lies between the collusive and the sequential-move follower outcomes; Region III corresponds to cases in which profit lies between the sequential-move follower outcome and the sequential-move leader outcome; Region IV corresponds to cases in which profit lies between the sequential-move leader outcome and Bertrand benchmarks. Dashed lines show additional isoprofit curves, each indicating increments equal to one-fourth of the difference between the collusive profits and the sequential follower profits. Assumes $d = 1$ under the linear demand system given by equation (22).

profits. Thus the algorithmic firm is the main beneficiary of coercion: it can appropriate a disproportionate share of the additional industry surplus generated by higher prices.

We now formalize general properties of coercive equilibrium. Motivated by the comparative statics in Figure 5, we now show that, under mild conditions ensuring that the within-period update is a punishment for the rival, firm a's equilibrium profit is nondecreasing in the speed advantage α. Firm a's equilibrium profit is also nondecreasing in the effective commitment $\beta\gamma$.

Proposition 3 (Profit Monotonicity in Speed)**.** *For the algorithmic firm, coercive equilibrium profit is weakly increasing in the speed advantage α if the algorithm's update acts as a punishment for deviations from the target price (i.e., $\pi_b(\hat{p}_b, \sigma(\hat{p}_b)) \leq \pi_b(\hat{p}_b, p_a^\dagger)$ for all deviations $\hat{p}_b \neq p_b^\dagger$).*

Proposition 4 (Profit Monotonicity in Commitment)**.** *For the algorithmic firm, coercive equilibrium profit is weakly increasing in commitment $\beta\gamma$.*

Intuitively, a higher α strengthens the effectiveness of rapid undercutting following a deviation and thereby allows the algorithmic firm to sustain more demanding target prices for the rival. The additional condition simply ensures that the algorithm's reaction is indeed a punishment relative to the pre-update price. If the algorithm were to respond by raising prices or otherwise increasing the rival's payoff, a faster reaction would make deviations more profitable, tightening the incentive constraint. In standard differentiated products settings with strategic complements, this condition is typically satisfied: punishments usually involve price cuts, and since the rival's profit is increasing in the algorithmic firm's price, facing a lower price $R_a(\hat{p}_b) < p_a^\dagger$ strictly reduces the rival's deviation profit.

The result for commitment holds generally and does not require the punishment condition. A higher level of commitment $\beta\gamma$ reduces the temptation to cut own price in the current period, because future profits under the algorithm become more important. Taken together, speed and commitment expand the set of credible punishments and make coercion more effective.

As an immediate corollary, the algorithmic firm's profit is maximized at the point of *maximal coercion*, $(\alpha, \beta\gamma) = (1, 1)$ under the conditions of Propositions 3–4. In that case the incentive compatibility constraint for the rival is as weak as possible, and the algorithmic firm can choose the most profitable target prices subject only to this single constraint.

The next result provides general conditions under which the algorithmic firm can secure more than its collusive profits.

Proposition 5 (Supracollusive Profits)**.** *Let* (p_a^C, p_b^C) *denote the joint profit–maximizing (collusive) prices, and let* (p_a^S, p_b^S) *denote the one-shot Stackelberg outcome in which firm* b *moves first and firm* a *best responds. If* $\pi_b(p_b^C, p_a^C) > \pi_b(p_b^S, p_a^S)$*, then there exist thresholds* $\overline{\alpha} \in (0, 1)$ *and* $\overline{\delta} \in (0, 1)$ *(with* $\delta \equiv \beta\gamma$*) such that for all* $\alpha > \overline{\alpha}$ *and* $\beta\gamma > \overline{\delta}$*, firm* a*'s coercive equilibrium per-period profit is strictly greater than* $\pi_a(p_a^C, p_b^C)$*.*

The condition on firm b is mild and is satisfied in standard demand systems. If joint-profit-maximizing prices give the slower firm more than it would receive as a static Stackelberg leader, then when α and $\beta\gamma$ are large enough, the collusive price vector satisfies firm b's incentive compatibility constraint. Once that constraint is slack, the algorithmic firm can adjust the target prices (lowering its own price or raising the rival's) to increase its profits above the collusive share while still deterring deviations. In this sense, coercion allows the algorithmic firm to use the threat of rapid undercutting as a commitment device to extract surplus from the rival.

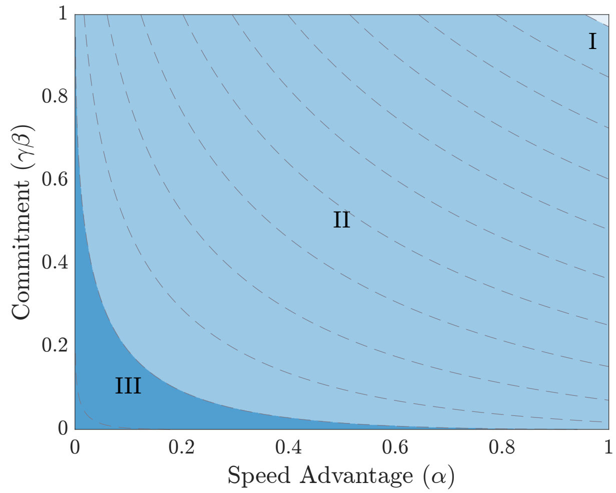
Though we have sharp predictions for profits, establishing comparative statics for prices requires additional restrictions about demand. To show monotonicity of prices, a sufficient condition is that demand is linear as in our numerical example from Section 3.5. Demand systems with different curvature properties may not maintain monotonicity in prices.

Assumption D1 (Linear Demand)**.** *Demand follows the linear specification in equation* (22)*.*

Proposition 6 (Price Monotonicity in Speed and Commitment)**.** *Suppose Assumptions A1, A2, and D1 hold, so that demand is given by the linear system in equation* (22)*. Then, both firms' equilibrium target prices* $(p_a^\dagger, p_b^\dagger)$ *are weakly increasing in the speed advantage* α *for each fixed* $\beta\gamma \in (0, 1)$*, and weakly increasing in the effective commitment parameter* $\beta\gamma$ *for each fixed* $\alpha \in [0, 1]$.

Key conditions for this to hold are that the binding deviation price $\hat{p}_b$ for the rival is strictly below the target price $p_b^\dagger$ and that the rival earns strictly lower profits facing the algorithmic

Figure 6: Consumer Surplus Regions by Pricing Speed and Commitment



Notes: The figure reports consumer surplus regions as a function of the algorithmic firm's speed advantage α and commitment $\beta\gamma$. Region I consists of parameter pairs for which consumer surplus falls below its level under joint-profit maximization (collusion); Region II corresponds to cases in which consumer surplus lies between the collusive and the sequential-move outcomes; Region III corresponds to cases in which consumer surplus lies between the sequential-move and Bertrand benchmarks. Dashed contour lines mark increments of one-tenth of the difference between consumer surplus under collusion and under the sequential-move equilibrium. Demand follows the linear specification in equation (22) with $d = 1$.

firm's best response than facing the target price: $\pi_b(p_b, R_a(p_b)) < \pi_b(p_b, p_a^\dagger)$. Assumption D1 is a convenient sufficient assumption under which the two conditions can be verified in closed form. Other well-behaved demand systems may also satisfy these inequalities. Under these conditions, a faster algorithm reduces the rival's payoff from deviating, because the punishment following any undercut arrives more quickly and more often. This relaxes the rival's incentive compatibility constraint and enables the algorithmic firm to sustain higher target prices while still deterring deviations.

4.2 Consumer Surplus

We now study the implications for consumer surplus. We assume that the demand system is generated by a representative consumer with quasi-linear utility over the differentiated goods and a numeraire, and we interpret consumer surplus as the consumer's money-metric indirect utility.

Figure 6 partitions the $(\alpha, \beta\gamma)$–space into regions according to how the coercive equilibrium compares with standard benchmarks, using our numerical example. The regions show when consumer surplus lies (i) below the collusive level (Region I), (ii) between the collusive and sequential-move levels (Region II), or (iii) between the sequential-move and Bertrand levels (Region III). Consumers are never better off than under the Bertrand outcome: for all values of the speed advantage and commitment, consumer surplus in the coercive equilibrium is weakly below the Bertrand level. For most parameter combinations, the reduction is meaningful, with consumer surplus falling between the values obtained under collusion and the sequential-move outcome (Region II).

When both α and $\beta\gamma$ are close to one, the coercive equilibrium yields the lowest consumer surplus (Region I). In this part of the parameter space, consumer surplus lies strictly below the collusive benchmark, even though firms are not explicitly coordinating on prices. While the algorithmic firm can have higher profits than the joint-profit-maximizing benchmark, total industry profits are lower. We summarize these observations as follows.

Remark 3. *Consumer surplus and total welfare can be lower in the coercive equilibrium than under joint profit maximization.*

The same mechanism highlighted in Section 4.1 drives this welfare result. When both α and $\beta\gamma$ are large, the algorithmic firm can sustain a high target price and credibly threaten to undercut any deviation by the slower rival. The rival is then coerced into setting a price that can exceed the collusive level, while the algorithmic firm prices slightly below. These prices leave consumers strictly worse off than under symmetric collusion.

Figure 6 also illustrates how the two technological dimensions of the algorithm—within-period speed and across-period commitment—interact. When either α or $\beta\gamma$ is close to zero, the algorithmic firm has limited ability to discipline its rival and the loss in consumer surplus relative to the Bertrand benchmark is modest. As either parameter increases, holding the other fixed, consumer surplus declines.

With restrictions on the curvature of demand, we can establish that consumer surplus declines monotonically with speed and commitment. One sufficient condition is that demand is linear, following the monotonicity of prices established in Proposition 6. Since consumer surplus is decreasing in prices, the ability of the algorithmic firm to use speed or commitment to raise target prices leads to a monotonic decline in consumer surplus.

Proposition 7 (Consumer Surplus Monotonicity)**.** *Suppose Assumptions A1, A2, and D1 hold, so that demand is given by the linear system in equation* (22)*. In coercive equilibrium, consumer surplus is weakly decreasing in both α and $\beta\gamma$.*

The above analysis shows that the combination of faster pricing and (imperfect) commitment can have first-order consequences for consumers, even in environments where the rival is not forward-looking and cannot sustain collusion. In the next subsection, we show that this qualitative finding is robust to a number of alternative modeling assumptions.

4.3 Robustness to Alternative Assumptions

We now examine how the coercive equilibrium responds to several modifications of the baseline environment. In turn, we relax the assumption that the non-algorithmic firm is purely myopic, consider alternative punishment technologies for the algorithmic firm, examine alternative timing conventions at reprogramming dates, and extend the market structure to an oligopoly in which a single algorithmic firm faces multiple rivals without algorithms.

Forward-Looking Rival

Assumption A1 that firm b is short-sighted essentially requires the algorithmic firm to enforce compliance using only current-period incentives. As noted in Section 2, this is a more stringent requirement than standard collusion, which utilizes future punishment values to relax incentive constraints. Here, we demonstrate that the coercive equilibrium persists even when the rival is forward-looking.

Suppose instead that firm b maximizes the sum of discounted profits. There exists an equilibrium strategy where neither firm b nor firm a conditions on the history of play. The optimal strategy of this form has firm b solving the following dynamic problem:

$$V_b(t, \mathcal{A}_\tau) = \max_{p_b}(1-\alpha)\pi_b(p_b, \rho_\tau) + \alpha\pi_b(p_b, \sigma_\tau(p_b)) + \hat{\beta}\gamma V_b(t+1, \mathcal{A}_\tau) + \hat{\beta}(1-\gamma)V_b(t+1, \mathcal{A}_{t+1})$$

where $\hat{\beta}$ is firm b's discount factor and $\mathcal{A}_{t+1}$ denotes the algorithm chosen by firm a in period $t+1$. The evolution of the state to period $t+1$ is independent of the choice of p_b in period t. Thus, a policy of choosing p_b to maximize current-period profits is an equilibrium strategy and the coercive equilibrium we characterize persists when firm b is forward-looking.

While allowing for arbitrary history dependence reintroduces the standard multiplicity of repeated-game equilibria, such outcomes rely on intricate, contingent punishment strategies that are fragile in the presence of complexity costs (e.g., Gale and Sabourian, 2005). The Markovian strategy discussed in Section 2 avoids this complexity by treating the algorithm as a fixed constraint. Consequently, the coercive equilibrium we characterize remains the natural complexity-minimizing selection.

Alternative Punishment Functions

Assumption A2 specifies that, following any deviation by firm b, the algorithm punishes with the static best response $R_a(\cdot)$. This is a natural benchmark and a standard assumption in the literature about collusion, and it is sufficient to generate coercion. It is, however, not the only possibility, and in this subsection we explore how alternative punishments affect the equilibrium.

We first consider a harsher punishment in which firm a commits to setting price equal to marginal cost after any deviation by firm b (so $p_a = 0$ in our linear-demand example). Appendix Figure A-7 plots equilibrium prices and profits as a function of the algorithmic technology. The qualitative relationship with the speed advantage α and the degree of commitment $\beta\gamma$

is similar to that under best-response punishments. In particular, faster pricing and stronger commitment raise equilibrium prices and profits for both firms. Stronger punishment increases firm b's incentive to comply and relaxes the incentive constraint. This allows the algorithmic firm to coerce its rival into setting even higher prices, thereby earning strictly higher profits for any given level of commitment.

We next consider a milder punishment in which the algorithmic firm employs price matching, so that $\sigma_a(p_b) = p_b$ following a deviation from the target price. Price-matching guarantees are frequently observed in practice and may soften competition (see, for example, Salop, 1986). Under price matching, the comparative statics with respect to α and $\beta\gamma$ remain: equilibrium prices and profits increase with the speed advantage and the degree of commitment (Appendix Figure A-8). When firm a can react instantaneously and the two firms are symmetric, however, the unique coercive equilibrium coincides with the fully collusive outcome, regardless of the level of commitment. In this case, a deviation by firm b leads immediately to symmetric prices. Since firm b could always deviate to the collusive price and face a "punishment" of earning full collusion profits, the algorithmic firm must offer firm b at least full collusion profits in equilibrium.

Compared to price matching, best-response punishments deliver larger gains to the algorithmic firm when commitment is strong, since they allow it to earn more than its full collusion profits. By contrast, with weak commitment a more severe best-response punishment may be suboptimal for firm a, and a price-matching rule can yield higher profits. In Section 5 we study a further punishment technology in which the algorithmic firm uses a pricing rule that is linear in the rival's price; this specification is particularly relevant when firm b follows a naive learning rule rather than solving a fully specified dynamic program.

Alternative Timing Assumptions

The baseline timing in Section 2 assumes that, at a reprogramming date τ, firm a chooses the new algorithm $\mathcal{A}_\tau$ while firm b simultaneously sets $p_{b\tau}$. Thus, our equilibrium concept is a simultaneous-move Nash equilibrium in pure strategies. Appendix A.2 relaxes this modeling choice by considering four alternative move orders in the reprogramming period, including environments in which the algorithm must punish during the reprogramming period, environments with a one-period implementation lag in which within-period punishment is not available in the reprogramming period, environments in which the target price $p_b^\dagger$ is chosen before ρ and $p_{b\tau}$, and environments in which firm a chooses the full algorithm first and firm b follows.

Under the fastest within-period response and full commitment (i.e., $(\alpha, \beta\gamma) = (1, 1)$), the maximal-coercion outcome characterized by (21) remains an equilibrium across all timing variants (Appendix Table A-1). Thus, the coercive outcome is not an artifact of the particular ordering assumed at reprogramming dates. Speed (α) continues to act as an enforcement technology (enabling within-period discipline of deviations), while commitment can be supplied either by persistence of the algorithm (through $\beta\gamma$ in the baseline) or, in some timing variants, by an additional first-mover commitment channel embedded in the reprogramming-period move order. Consistent with this interpretation, timing that gives firm a an explicit commitment advantage in the reprogramming period can expand the set of environments in which coercion is sustainable, whereas timing that temporarily disables punishment in the reprogramming period shrinks it. Overall, the models with alternative timing reinforce the interpretation of supracompetitive outcomes as arising from the interaction of speed and commitment, while confirming that the maximal coercion outcome is robust to the details of the within-period timing.

N-firm Oligopoly

Finally, we extend the analysis to an oligopoly in which a single algorithmic firm competes against several rivals that do not use pricing algorithms. Appendix Section A.3 sets out the formal environment. The algorithmic firm again chooses a target price vector and an associated algorithm, now subject to the incentive compatibility constraints of all non-algorithmic rivals. In the case with two non-algorithmic rivals, the comparative statics with respect to the speed advantage and commitment closely resemble those in Figure 3 and Figure 4. While higher α and higher $\beta\gamma$ still raise equilibrium prices and profits, the additional incentive constraints limit the magnitude of the effects.

Appendix Figure A-2 reports the maximal coercion equilibrium as the number of non-algorithmic rivals increases, under a generalization of the linear demand specification in equation (22). As the market becomes more crowded, both the algorithmic firm's price and its rivals' prices decline monotonically, and profits fall. The erosion of market power is gradual, however: even with ten non-algorithmic rivals, the share-weighted average price remains roughly 18 percent above the Bertrand benchmark in our example. Thus, while additional competitors discipline prices, the combination of a speed advantage and (imperfect) commitment still allows a single algorithmic firm to sustain substantially supracompetitive outcomes.

5 Learning by a Naive Rival

The analysis so far has taken the rival without an algorithm to be fully aware of the algorithmic firm's pricing rule and of the associated profit consequences. In particular, when firm a commits to a punishment strategy, firm b is assumed to understand how the within-period response function $\sigma(\cdot)$ maps its current price into payoffs and therefore to best respond period by period. It is natural to ask how robust our coercion results are when this informational assumption is relaxed.

In many markets it is unlikely that rivals perfectly decode the algorithm used by a competitor, or even that they model the interaction as a dynamic game. Instead, managers may run pricing experiments and adjust prices myopically in directions that appear to raise current profits. In this section we therefore treat firm b as "naive": it does not know the algorithmic firm's strategy or the structure of demand, and it learns only from the history of its own prices and realized profits.

Two questions then arise. First, can the algorithmic firm still induce supracompetitive prices when the rival behaves in this simple, data-driven way? Second, if so, how quickly do prices converge and how sensitive is convergence to demand shocks? To address these questions we introduce a learning rule for the slow firm and study its interaction with a class of linear pricing algorithms. We show that the learning dynamics converge to a stationary coercive outcome, and that convergence is rapid even when demand is noisy.

5.1 Gradient Learning and Linear Punishment Strategies

It is often more plausible to view market outcomes as the long-run result of adaptive behavior than as the product of fully rational, forward-looking play. A large literature therefore studies learning in games (see, for example, Fudenberg and Levine, 2016). In this spirit, we model firm b as experimenting with prices and updating them in response to observed profits, rather than solving explicitly for the equilibrium of the dynamic pricing game. This assumption is consistent with evidence that many online retailers run pricing experiments while treating competitors' behavior as exogenous (e.g., "A/B testing").[15]

A variety of learning rules have been proposed, including fictitious play, reinforcement learning, and gradient learning. Different learning models can have different convergence properties. The recent literature has shown that different classes of learning algorithms may or may not converge to Bertrand–Nash equilibria.[16]

In our setting we focus on *gradient learning* by the slower firm. Gradient learning is a particularly naive rule that requires very little information. Firm b does not condition on the rival's prices, does not attempt to infer the pricing rule used by firm a, and does not estimate a demand system. Instead, it updates its price in the direction of the estimated gradient of its own profit function, thereby capturing the idea that managers increase price when the last increase was profitable and decrease price otherwise.

Formally, the slow firm starts with two initial price choices, p_b^0 and p_b^1 with $p_b^0 \neq p_b^1$, and uses the change in realized profits between these two experiments to approximate the slope of its profit function. For period $t \geq 1$ the update rule is

$$p_{b,t+1} = p_{bt} + \lambda \, \widehat{\frac{\partial \pi_b}{\partial p_b}}\Big|_{p_{bt}}, \tag{24}$$

[15]For instance, given the complexity of demand and the number of products, managers may test discrete price changes and adopt changes that appear to raise profits, without explicitly modeling rivals' reactions. See discussion in Hansen et al. (2021).

[16]For example, the Q-learning algorithms studied in Calvano et al. (2020) do not generally converge to Bertrand–Nash, while Asker et al. (2024) show that alternative assumptions on this class of algorithms do lead to convergence.

where $\lambda > 0$ is a step-size parameter and $\widehat{\partial\pi_b/\partial p_b}|_{p_{bt}}$ denotes firm b's estimate of the gradient of its current-period profit with respect to its own price. Throughout we assume that λ is small enough to ensure that the sequence $\{p_{bt}\}$ is stable. Under these assumptions, a simultaneous pricing game in which both firms use gradient learning converges to the Bertrand–Nash equilibrium (Anufriev et al., 2013).

The presence of learning on the rival side also changes the algorithmic firm's problem. To keep the analysis tractable, and to capture pricing rules commonly observed in practice, we restrict attention to linear punishment strategies. In contrast to the discontinuous trigger rule used in Assumption A2, we now suppose that when firm a targets a price pair $(p_a^\dagger, p_b^\dagger)$ it sets $p_a = p_a^\dagger$ at the beginning of the period and subsequently adjusts its price linearly in firm b's current price according to

$$\sigma(p_b) = \begin{cases} p_a^\dagger - \phi\big(p_b^\dagger - p_b\big) & \text{if } p_a^\dagger - \phi\big(p_b^\dagger - p_b\big) \geq 0, \\ 0 & \text{otherwise.} \end{cases} \tag{25}$$

Without loss of generality, we normalize marginal costs to 0 for this section, such that $p_a = 0$ corresponds to marginal-cost pricing.

The punishment depends on the deviation $p_b - p_b^\dagger$ from the target price, and the intensity of the response is governed by the parameter $\phi > 0$. By construction $\sigma(p_b^\dagger) = p_a^\dagger$. We focus on linear rules because, as we show below, they interact particularly simply with gradient learning and ensure convergence of the naive firm's experimentation to the algorithmic firm's desired target. Moreover, linear dependence on rivals' prices, such as rules that undercut competitors' observed prices by a fixed amount or percentage, are commonly observed in online marketplaces (Chen et al., 2016; Musolff, 2024).

In terms of implementation, a coercive linear rule can be thought of as a very simple pricing heuristic. A fast repricing tool can be instructed to undercut aggressively whenever the rival prices below the target prices. Over time, the rival observes that choosing prices below a given level lead to low profits, and its strategy converges to choosing the price targeted by the coercive algorithm, as we illustrate below. Intuitively, this linear rule operates as a tax on the rival's price reductions. By committing to lower its price proportionally whenever the rival cuts price, the algorithm effectively rotates the rival's residual demand curve, making it appear more inelastic to the rival and thereby diminishing the incentive to undercut.

Throughout this section we maintain the assumption that the algorithmic firm can fully commit to the linear pricing rule in (25). In particular, the firm does not adjust its algorithm over time in order to manipulate the rival's learning speed; instead, it chooses a rule once and then allows the learning dynamics of firm b to play out against this fixed environment.

5.2 Equilibrium Target Prices

We now analyze the case in which the algorithmic firm commits to a linear pricing rule while the rival uses simple gradient learning. The algorithmic firm chooses a target price vector $(p_a^\dagger, p_b^\dagger)$ and commits to an update rule that adjusts its price proportionally to the rival's current choice. To pin down the slope of this rule, we restrict attention to linear pricing rules that pass through the origin and the target point $(p_b^\dagger, p_a^\dagger)$. For any non-negative price chosen by the slower rival, the faster firm then responds along this line, so that the slope is $\phi = p_a^\dagger / p_b^\dagger$. Throughout this subsection we maintain $p_b \geq 0$.

Under these assumptions, the objective function can be written as:

$$\begin{aligned} &\max_{(p_a^\dagger, p_b^\dagger)} \pi_a(p_a^\dagger, p_b^\dagger) \\ &\text{s.t.} \quad p_b^\dagger = \operatorname*{argmax}_{p_b} (1-\alpha)\pi_b(p_b, p_a^\dagger) + \alpha \pi_b(p_b, \phi p_b) \end{aligned} \tag{26}$$

where the initial price in each period, ρ, is equal to $p_a^\dagger$, and where the update function is given by $\sigma(p_b) = \phi p_b$. As in the previous subsection, the algorithmic firm chooses a target price vector subject to the incentive compatibility constraint of the slower rival.

The slower rival's first-order condition associated with the inner maximization problem is

$$(1-\alpha)\frac{\partial \pi_b(p_b, p_a^\dagger)}{\partial p_b} + \alpha \frac{\partial \pi_b(p_b, \phi p_b)}{\partial p_b} + \alpha\phi \frac{\partial \pi_b(p_b, \phi p_b)}{\partial p_a} = 0 \tag{27}$$

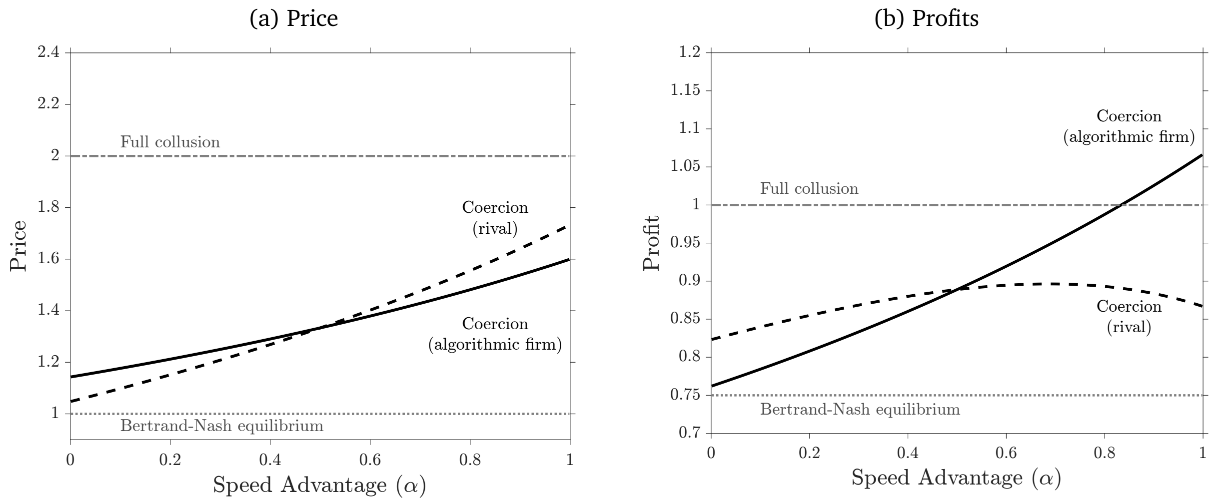
Because $\phi = p_a^\dagger / p_b^\dagger$, this condition implicitly defines the slower firm's best-response function $p_b^*(p_a^\dagger)$. Similarly, let $p_a^*(p_b^\dagger)$ denote the optimal price for the algorithmic firm as a function of the target $p_b^\dagger$. Substituting these implicit response functions into the objective, the algorithmic firm's first-order condition is given by:

$$\frac{\partial \pi_a(p_a^*(p_b^\dagger), p_b^\dagger)}{\partial p_a} \frac{\partial p_a^*(p_b^\dagger)}{\partial p_b^\dagger} + \frac{\partial \pi_a(p_a^*(p_b^\dagger), p_b^\dagger)}{\partial p_b} = 0 \tag{28}$$

With the constraint $p_b^\dagger = p_b^*(p_a^\dagger)$, these conditions characterize the long-run coercive equilibrium when the algorithmic firm is constrained to use a linear pricing rule and the rival follows gradient learning.

Figure 7 illustrates the long-run equilibrium for different values of the algorithmic firm's speed advantage. Panel (a) shows that prices for both firms lie strictly above the Bertrand benchmark and rise with the speed advantage α. When the algorithmic firm is sufficiently fast, it sets the lower of the two prices, using its rapid punishment to induce the slower rival to maintain a higher price. The qualitative pattern is similar to that obtained with the trigger strategy in Section 3, but the linear restriction on the pricing rule limits the extent to which the

Figure 7: Prices and Profits with Linear Pricing Rule, by Algorithmic Firm's Pricing Speed



Notes: Panel (a) reports equilibrium prices and panel (b) reports equilibrium profits under Bertrand competition with simultaneous pricing, under the coercive equilibrium when the faster firm uses a linear punishment rule and has full commitment ($\beta\gamma = 1$), and under joint profit maximization. In the coercion case, firms choose prices simultaneously at the start of the period and the algorithmic firm then adjusts its price within the period with speed advantage α. The simulations assume $d = 1$ under the linear demand system in equation (22).

rival's price can be pushed above the collusive level.

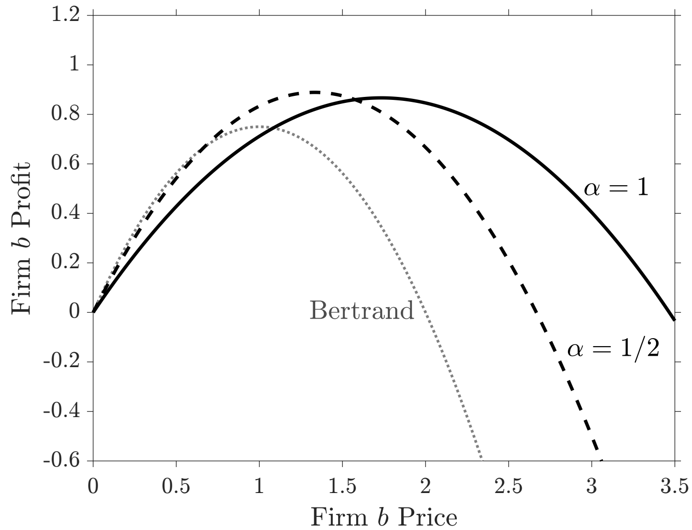
Panel (b) displays the corresponding profits. The algorithmic firm's profit is increasing in α, while the slower rival's profit is non-monotone. Even with the linear restriction the algorithmic firm can obtain more than its full collusion profits when its speed advantage is large. At $\alpha = 1$, the rival earns higher profits than under the coercive nonlinear strategy examined earlier (Figure 3), reflecting the fact that the linear pricing rule provides a weaker punishment for deviations.

5.3 Convergence Dynamics and Robustness

The linear pricing rules studied above imply that, from the perspective of the slow firm, the mapping from its own price to profit is smooth and single-peaked. In particular, when firm a adopts the optimal linear rule, firm b faces a one-dimensional choice problem with a concave objective. This structure is important for learning: it means that even a very simple adjustment rule, which ignores the opponent's strategy and treats the environment as essentially static, will tend to push prices toward the coercive target. In this subsection we formalize this observation and illustrate it with simulations.

Proposition 8 (Convergence of Learning Dynamics)**.** *Under linear demand, when firm a implements the optimal linear pricing rule, firm b's profits are increasing in price for any $p_b < p_b^\dagger$ and decreasing in price for any $p_b > p_b^\dagger$. Therefore, provided λ is not too large, the use of gradient*

Figure 8: Perceived Profit Function for the Naive Rival



Notes: The figure shows perceived profits for firm b as a function of its own price, under different values of the algorithmic firm's speed advantage α. The dotted line plots the benchmark case with simultaneous pricing in which the rival sets the Bertrand–Nash price. The other lines show the perceived profit function when firm a uses the optimal linear rule and has a speed advantage $\alpha = 0.5$ or $\alpha = 1$. In each case, perceived profit is single-peaked, but the maximizer is shifted to the right as α increases. The example assumes $d = 1$ under the linear demand system given by equation (22).

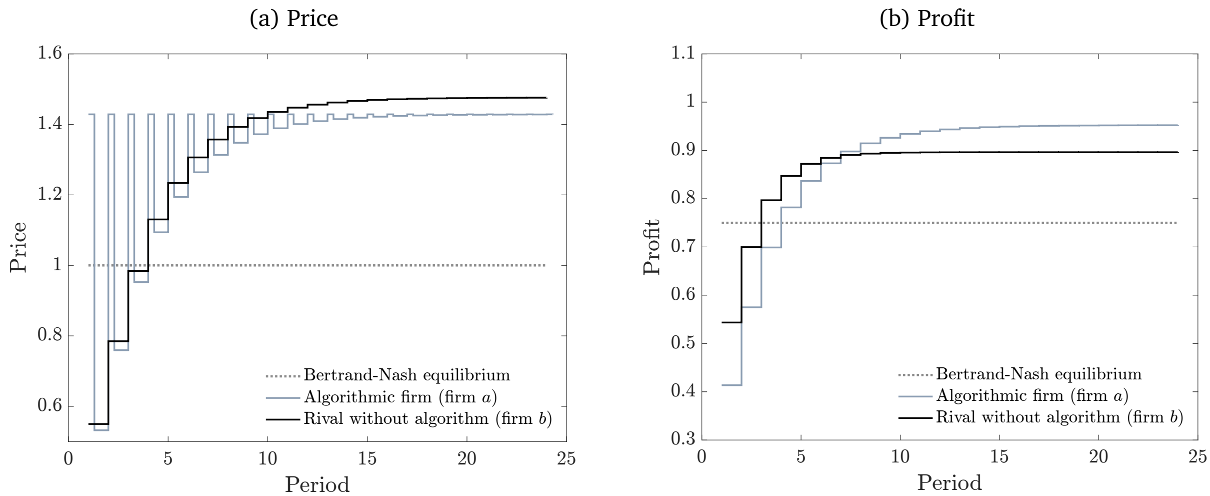
learning by firm b will result in convergence to the target price chosen by firm a for any initial price pair $(\hat{p}_b^0, \hat{p}_b^1)$.

The linear pricing rule implemented by firm a effectively transforms firm b's perceived profit function by creating a synthetic residual demand curve that is steeper below the target price. The proposition above implies that this perceived profit function is a strictly concave function in firm b's own price. As a result, the coercive target $p_b^\dagger$ is the unique maximizer of the perceived profit, and standard gradient learning converges to this point as long as λ is sufficiently small.

Figure 8 illustrates the perceived profit function faced by the naive rival. Under simultaneous pricing when firm a sets the Bertrand price, $p_a = 1$, firm b's profit is also maximized at $p_b = 1$. When firm a has a speed advantage and uses the linear pricing rule, the shape of firm b's perceived profit function is roughly preserved, but its peak shifts to higher prices. For instance, in the example with $d = 1$, the profit-maximizing price for firm b moves to 1.33 when $\alpha = 0.5$ and to 1.73 when $\alpha = 1$. Thus, from the perspective of firm b, it is optimal to raise price above the competitive level, even though the higher profitability is entirely induced by the algorithmic firm's rapid response.

In Figure 9, we simulate an example in which firm a implements the linear pricing rule and has a moderate speed advantage ($\alpha = 0.7$). Firm b is naive and starts by setting an arbitrary price and attempts to maximize profits using gradient learning. We assume firm b approximates the derivative of profits using the most recent two prices and profit realizations. Therefore, after

Figure 9: Simulated Learning with Linear Pricing Rule



Notes: The figure shows simulated price paths when the naive firm is assumed to use gradient learning. The black line shows price and profit for firm b, where price is updated in period $t+1$ following $\hat{p}_b^{t+1} = \hat{p}_b^t + \lambda(\pi^*_{bt} - \pi^*_{b,t-1})/(\hat{p}_{bt} - \hat{p}_{b,t-1})$ with a price adjustment step size, λ, of 0.2. The blue line shows price and profit for firm a, which uses a linear pricing rule described in Section 5.2 with pricing speed given by $\alpha = 0.7$. The dashed lines show the price and profits under the Bertrand benchmark. Assumes $d = 1$ under the linear demand system given by equation (22).

the firm sets price $\hat{p}_b^t$ at time t, the firm's estimate of the derivative of its profit with respect to price are given by

$$\widehat{\frac{\partial \pi_b}{\partial p_b}}\bigg|_{\hat{p}_{bt}} = \frac{\pi^*_{bt} - \pi^*_{b,t-1}}{\hat{p}_{bt} - \hat{p}_{b,t-1}} \tag{29}$$

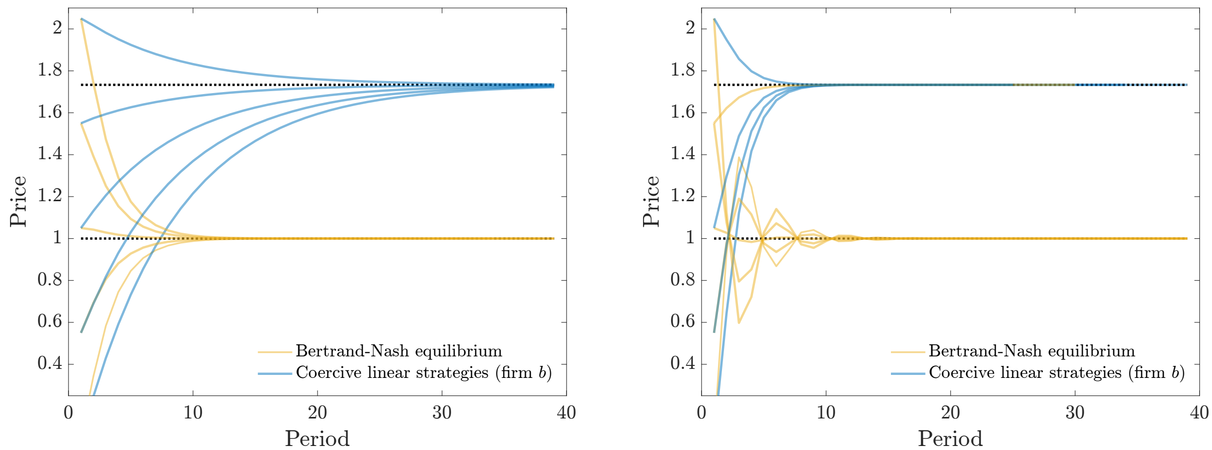
where π^*_{bt} is the realized profit for period t. Note that firm b need not observe firm a's price but only its own realized profits over the period. Then, the firm updates price according to equation (24). We assign a value of $\lambda = 0.2$.

For the simulation in Figure 9, we start from an initial price below the Bertrand level. At each step, firm b computes the finite-difference gradient in (29) and adjusts its price accordingly. Panel (a) shows that firm b steadily raises its price, while firm a responds via the linear pricing rule derived in Section 5.2. Panel (b) shows the corresponding profit paths. As long as the perceived marginal profit is positive, the naive firm continues to increase price. Because the punishment encoded in the algorithm becomes less severe at higher prices, the gradient remains positive until the coercive target $p_b^\dagger$ is reached. At this target price, the system reaches a fixed point, and the algorithmic firm's initial price in each period coincides with the punishment price implied by the linear rule.

Figure 10 examines convergence from multiple starting prices when $\alpha = 1$. First, we consider the standard case of a simultaneous pricing game as a benchmark for comparison. As is well known, prices converge to the Bertrand equilibrium from a variety of starting values in

Figure 10: Convergence of Simulated Learning with Linear Pricing Rule versus Simultaneous Pricing

(a) Low Price Adjustment Step Size ($\lambda = 0.2$)   (b) High Price Adjustment Step Size ($\lambda = 0.6$)



Notes: The figure shows simulated price paths for firm b with different initial prices when firm b is assumed to use gradient learning. The orange lines show a simultaneous pricing game that results in Bertrand prices. The blue lines show the game in which the algorithmic firm has maximum pricing speed ($\alpha = 1$) and uses the linear pricing rule described in Section 5.2. Panel (a) shows learning with price adjustment step size of 0.2 and panel (b) shows learning with a price adjustment step size of 0.6. Assumes $d = 1$ under the linear demand system given by equation (22).

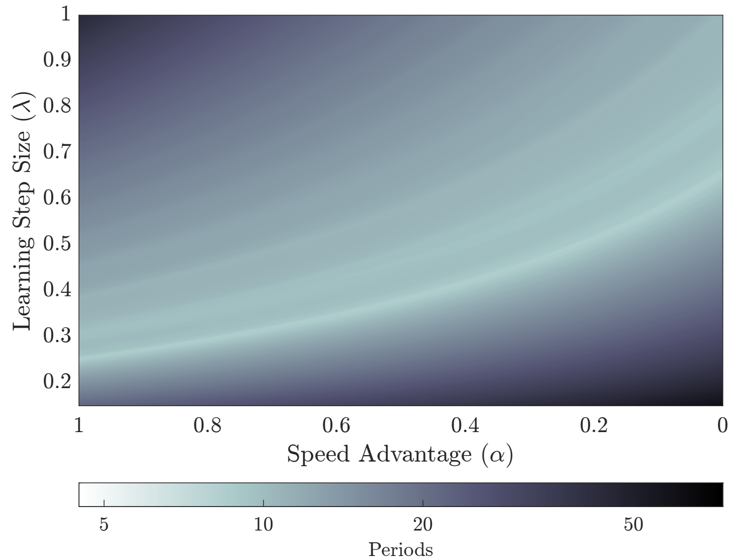
a standard simultaneous pricing game under gradient learning. This is shown by the orange lines. Convergence takes about 15 periods as seen in both panel (a) with a low step size and panel (b) with a higher step size. Under coercive linear strategies, the naive firm's price converges to a supracompetitive level for a variety of initial values, as illustrated by the blue lines in Figure 10. Comparing panel (a) and panel (b), convergence is somewhat slower than the standard simultaneous pricing case for a low price adjustment step size but is somewhat faster for a high price adjustment step size.

Regardless of the initial price, convergence occurs quickly, though it does depend on pricing speed and the choice of step size, λ. For each value of α and λ, we simulate the price paths for a grid of 1,000 initial prices. Figure 11 shows the average number of periods to convergence to within 0.1 percent of the target prices for a range of α and λ.[17] For all values of the learning parameter less than one, the number of periods to convergence is reasonable in our simulations. Moreover, for every value of α, there is a learning parameter λ such that the coercive strategies converge to the target prices in less than 10 periods on average. Higher step sizes tend to converge faster when α is large. Consistent with Proposition 8, the simulations always converge to the coercive equilibrium.[18]

[17]To ensure that prices have converged, we identify the first instance for which prices fall within a given tolerance and stay within that tolerance for the next 10 periods in our simulations. Given the nature of gradient learning, this provides confidence that the price does not drift outside of this tolerance in future periods.

[18]Appendix Table A-2 provides statistics on the number of periods to convergence under more and less stringent

Figure 11: Convergence Speed by Pricing Speed and Learning Step Size



Notes: The figure shows the number of periods until convergence for different simulation parameters when a naive firm is assumed to use gradient learning. The x-axis is the speed advantage of the algorithmic firm (α) and the y-axis is the learning step size (λ). For each value of α and λ, we simulate price paths from 1,000 values of initial starting prices from a grid between 0 and 2 and then average the number of periods until convergence. Convergence is defined as being within 0.1% of the equilibrium price. Darker color indicates a larger number of periods until convergence. Simulations assume the algorithmic firm uses the linear pricing rule described in Section 5.2. Assumes $d = 1$ under the linear demand system given by equation (22).

We next examine the robustness of these dynamics to unobserved demand shocks. Suppose the linear demand system in equation (22) is modified to

$$D_i(p_i, p_{-i}) = x_t - \left(\frac{1}{4} + \frac{d}{2}\right) p_i + \frac{d}{2} p_{-i}, \tag{30}$$

where the shock x_t is drawn independently from a uniform distribution with $E[x_t] = 1$. In each period, firm b observes only its own prices and realized profits and continues to update prices using the same gradient rule. The presence of shocks introduces noise into the finite-difference estimate in (29), so the firm occasionally moves in a direction that is locally unprofitable. Nonetheless, across many simulated paths, prices converge quickly to a distribution centered around the coercive target, and remain close to the target values when demand variation is modest. We illustrate convergence with demand shocks in Appendix Figure A-10.

Taken together, these results highlight that the coercive equilibrium is robust. When an algo-

convergence criteria. When the convergence criterion is within 1 percent of the target price, λ can be chosen such that convergence takes no longer than 9 periods for a wide range of starting values.

Appendix Figure A-9 shows the mean number of periods until convergence for $\alpha = 1$ compared to the simultaneous pricing case for different learning parameters. The fact that convergence in the coercive linear strategy case is faster for large step sizes is due in part to the fact that price oscillations are less likely during the learning process, as seen in panel (b) of Figure 10.

rithmic firm uses a simple linear rule with a speed advantage, even a very naive rival that relies on myopic gradient learning and observes only its own profits will be driven toward supracompetitive prices. Convergence is typically fast, and remains so in the presence of demand shocks. This stands in contrast to more complex reinforcement-learning algorithms, which often require many thousands of iterations to approach a stable outcome.[19]

We consider relatively simple pricing algorithms that are linear in a rival's price. Pricing algorithms could in principle take more general forms, allowing greater flexibility in strategically modifying a rival's perceived profit function. For instance, a pricing algorithm that is a nonlinear function of rivals' prices could potentially coerce rivals using naive learning strategies to raise prices to even higher levels than those under linear pricing rules. Nonlinear algorithms could potentially ensure that a naive firm is maximizing a concave profit function for more general demand systems, implying that there is convergence to supracompetitive prices under a variety of naive learning strategies.

6 Implications for Platform Design

Online retail platforms are often able to determine the pricing tools available to different sellers that compete on their platform. For example, a platform can determine (i) the speed with which sellers can observe and respond to rivals' prices (e.g., through API latency, batching, or minimum time-between-change rules), and (ii) the extent to which sellers can commit to pricing rules that persist over time (e.g., through restrictions on re-optimization, monitoring, or availability of proprietary tools). These choices are particularly salient when a platform is vertically integrated and sells its own products alongside third-party sellers. In this case, the platform may effectively grant itself different pricing capabilities than it grants to outside sellers.

In this section, we endogenize the algorithmic technology. We view the speed advantage and the commitment parameter not as exogenous primitives, but as design variables chosen by a platform seeking to maximize a weighted objective function. This simple framework allows us to analyze how platform incentives—specifically the degree to which they internalize consumer welfare versus seller profits—drive the adoption of coercive pricing technologies.

6.1 Platform Objective

Assume two sellers, a and b, compete on a platform in the baseline coercive pricing game. Seller a may be equipped with a pricing algorithm characterized by a speed advantage $\alpha \in [0, 1]$ and a commitment parameter $\gamma \in [0, 1]$. For simplicity, we assume $\beta \to 1$ in this section, which would

[19]Q-learning algorithms can take hundreds of thousands of iterations to converge; in some settings, it may be possible to resolve this with offline learning (Calvano et al., 2020).

be the case when periods are short.[20] Seller b does not use an algorithm and instead chooses prices to maximize per-period profits as in Assumption A1, while seller a uses best-response punishments as in Assumption A2.

We focus on the symmetric linear demand system in equation (22). For any (α, γ), let $\pi_i^*(\alpha, \gamma)$ denote seller i's per-period profit and let $CS^*(\alpha, \gamma)$ denote per-period consumer surplus in the (unique) coercive equilibrium induced by (α, γ).

The platform internalizes seller profits to different degrees. Let $\kappa_a, \kappa_b \geq 0$ capture the extent to which platform revenue increases with each seller's profit (e.g., through commissions, preferred seller arrangements, or vertical integration). The platform may also value consumer surplus, for instance because demand for the platform is increasing in the consumer experience. Let $\kappa_{CS} \geq 0$ denote the platform's weight on consumer surplus.

The platform chooses the algorithmic technology (α, γ) to maximize

$$\max_{\alpha, \gamma \in [0,1]} \ \kappa_{CS}\, CS^*(\alpha, \gamma) \ + \ \kappa_a\, \pi_a^*(\alpha, \gamma) \ + \ \kappa_b\, \pi_b^*(\alpha, \gamma). \tag{31}$$

Two comparative statics from earlier results discipline this problem. First, under linear demand (Assumption D1), consumer surplus is weakly decreasing in both speed and commitment (Proposition 7). Second, the algorithmic seller's profit is increasing in both speed and commitment (Propositions 4–3). Together, these imply a natural tension: higher (α, γ) benefits seller a and harms consumers, while the effect on seller b (and thus on the term $\kappa_b \pi_b^*$) can be ambiguous.

6.2 Platform Incentives

We begin with a welfare benchmark in which the platform values consumer surplus and seller profits symmetrically.

Proposition 9 (Welfare Benchmark)**.** *Suppose demand is given by equation* (22) *and maintain Assumptions A1–A2. If the platform weights consumer surplus and each seller's profit equally,* $\kappa_{CS} = \kappa_a = \kappa_b > 0$*, then the platform optimally chooses* $(\alpha, \gamma) = (0, 0)$*. In this benchmark, the platform does not permit pricing technology that confers either a speed advantage or long-run commitment.*

Within the class of outcomes implementable by changing (α, γ) in our coercion model, higher speed and stronger commitment move equilibrium prices weakly upward (Proposition 6). Under quasi-linear preferences, higher prices reduce equilibrium quantities and thus reduce total surplus. Hence, among these implementable equilibria, welfare is maximized by

[20] A platform does not literally choose the sellers' discount factor β. However, it can affect the effective commitment $\beta\gamma$ by changing the period length (how often prices can change at baseline), by determining how easily pricing rules can be revised or re-optimized, or by restricting sellers to use rules that persist over time.

choosing the least permissive technology, $(\alpha, \gamma) = (0, 0)$, which collapses the game to the standard Bertrand benchmark.

Next, consider a platform that disproportionately internalizes seller a's profit—for example because it is vertically integrated with seller a or because seller a faces a higher effective take rate.

Proposition 10 (Preferential Treatment). *Fix $\kappa_{CS} \geq 0$ and $\kappa_b \geq 0$. There exists a threshold $\bar{\kappa}_a(\kappa_{CS}, \kappa_b)$ such that, if $\kappa_a > \bar{\kappa}_a(\kappa_{CS}, \kappa_b)$, any optimal platform policy satisfies $(\alpha^*, \gamma^*) \neq (0, 0)$. Moreover, as κ_a becomes large relative to (κ_{CS}, κ_b), the platform selects the maximal technology $(\alpha^*, \gamma^*) = (1, 1)$.*

This result follows from the fact that $\pi_a^*(\alpha, \gamma)$ is increasing in both arguments (Propositions 3–4). When the platform places sufficiently high marginal value on seller a's profit, it prefers to expand the set of credible punishments available to seller a: greater speed makes punishments more immediate within the period, and greater commitment makes them more persistent across periods. When these incentives dominate any weighted losses in consumer surplus (and any effect on seller b), the platform optimally moves away from $(0, 0)$; in the limit where the platform effectively maximizes seller a's profit alone, it chooses $(1, 1)$.

Finally, holding fixed the platform's weights on seller profits, lowering the platform's concern for consumers pushes the platform toward permitting speed and/or commitment.

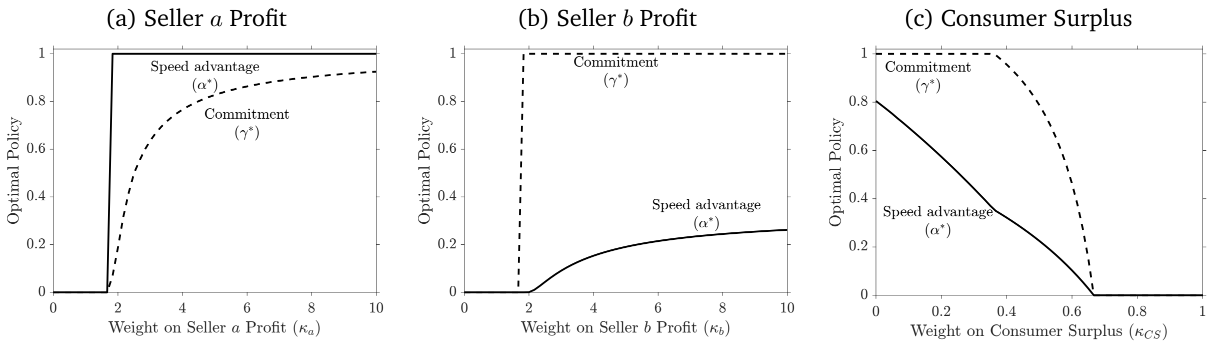
Proposition 11 (Consumer Surplus Threshold). *Fix (κ_a, κ_b) with $\kappa_a + \kappa_b > 0$. There exists a threshold $\bar{\kappa}_{CS}(\kappa_a, \kappa_b) > 0$ such that if $\kappa_{CS} \in [0, \bar{\kappa}_{CS}(\kappa_a, \kappa_b))$, any optimal platform policy satisfies $(\alpha^*, \gamma^*) \neq (0, 0)$. In particular, when the platform places sufficiently little weight on consumer surplus, it permits a positive speed advantage and/or commitment for seller a.*

Platforms need not explicitly coordinate prices to generate supracompetitive outcomes. It is enough that platform payoffs are sufficiently aligned with seller profits and sufficiently weakly aligned with consumer surplus; the platform then optimally chooses rules that make coercive pricing technology feasible.

Figure 12 illustrates how the platform uses the two instruments. We examine the platform's optimal choice of within-period speed advantage, α^* (solid), and the effective cross-period commitment, γ^* (dashed), as the platform varies one objective weight at a time around the normalization $\kappa_{CS} = \kappa_a = \kappa_b = 1$.

Panel (a) varies the weight on seller a's profit. As κ_a rises, the platform quickly moves to the corner $\alpha^* = 1$ and then increases effective commitment more gradually. Intuitively, speed is the enabling ingredient for coercion: without sufficiently fast within-period reactions, punishments arrive too late to strongly reshape seller b's static incentives. Once speed is high, additional commitment further increases seller a's ability to sustain more aggressive target prices. Conversely, as the weight on seller b's profits rises, the platform initially enables full

Figure 12: Optimal Platform Policy as Platform Objectives Vary



Notes: Each panel plots the platform's optimal choice of speed advantage α^* (solid) and effective commitment γ^* (dashed) solving (31) while isolating different components of the platform's objective. In each panel, the indicated weight varies while the remaining weights are held fixed at 1. Computed under (22) with $d = 1$.

commitment before gradually enabling a speed advantage since commitment alone benefits seller b. This can be seen in panel (b).

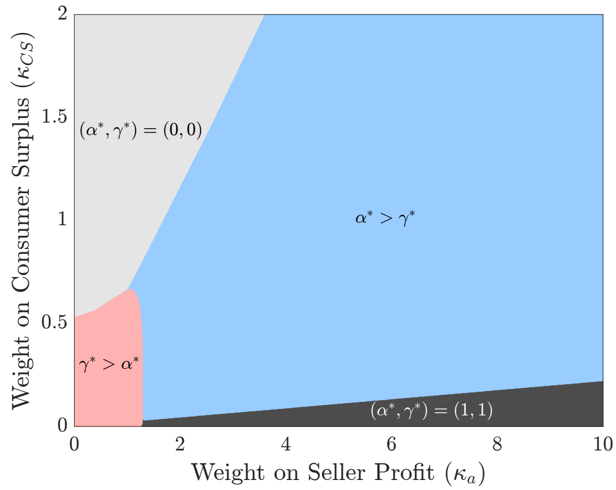
Panel (c) varies the weight on consumer surplus. As κ_{CS} rises, the platform first reduces the speed advantage while keeping effective commitment high, and then shuts down commitment as well; beyond a threshold it selects $(\alpha^*, \gamma^*) = (0, 0)$. In other words, once the platform places enough weight on consumer surplus, it is optimal to eliminate both technological features that expand the set of coercive equilibria.

Figure 13 summarizes the platform's optimal policy regimes as a function of κ_a and κ_{CS} holding $\kappa_b = 1$. The extreme regions match the comparative statics in Figure 12. When consumer surplus receives sufficiently high weight relative to seller profits (upper-left region), the platform selects $(\alpha^*, \gamma^*) = (0, 0)$, eliminating both speed advantage and commitment across periods. When the platform places sufficiently large weight on seller a's profit and relatively little on consumer surplus (lower-right region), it selects the maximal technology $(1, 1)$.

Between these extremes, Figure 13 shows two intermediate regimes that help interpret how the platform softens competition away from the corner solutions. In the blue region, $\alpha^* > \gamma^*$ and the platform prioritizes speed over commitment, allowing rapid within-period reactions while keeping long-run commitment weaker. In the pink region, $\gamma^* > \alpha^*$ and the platform prioritizes commitment over speed, permitting persistent pricing rules but constraining rapid reactions.

To summarize, platforms that behave more like welfare maximizers (or that place sufficiently high weight on consumer surplus relative to seller profits) prefer to eliminate speed advantages and persistent pricing rules. Platforms that internalize the profits of sellers—either because they place little weight on consumer surplus or because they explicitly favor a particular seller—have an incentive to enable the technological conditions that sustain coercive pricing equilibria.

Figure 13: Platform Policy Regions by Platform Objective



Notes: The figure reports policy regions as a function of the weight on consumer surplus κ_{CS} and seller a's profits κ_a. The weight on seller b's profits κ_b is held fixed at 1. Computed under equation (22) with $d = 1$.

These findings complement the literature focused on how platforms can facilitate or mitigate coordination through information sharing or demand steering (e.g., Johnson et al., 2023; Harrington, 2022). In addition to these channels, a platform can simply choose the pricing technology available to sellers—specifically speed and commitment—to either facilitate competition or enable supracompetitive prices through coercion.

7 Conclusion

We have analyzed the extent to which a single firm that delegates pricing to an algorithm can sustain supracompetitive outcomes. When the algorithm enables a speed advantage and commitment, prices can be substantially higher than competitive levels and consumer surplus can be lower than levels obtained with full collusion. This is true even when its rivals do not (and cannot) support collusion through forward-looking strategies. The key implication is that high prices need not reflect mutual coordination: they can arise from one firm's ability to make low-price deviations immediately unprofitable and to commit to doing so.

The mechanism is straightforward. Speed lets the algorithmic firm punish quickly so deviations by a rival deliver little or no gain, even if the rival is myopic. Commitment—through the persistence of the deployed pricing rule—then makes this threat optimal ex ante for the algorithmic firm. Together these features expand the set of price pairs the algorithmic firm can implement relative to standard leader–follower benchmarks. Importantly, the channel is distinct from repeated-game collusion and does not require symmetry, mutual patience, or sophisticated strategic reasoning by all firms.

We showed that this coercive mechanism is robust to several modifications of the environment. Allowing the rival to be forward-looking leaves our equilibrium intact. Alternative punishment technologies, such as instantaneous price matching, can support the fully collusive outcome as a unique equilibrium when the speed advantage is large, but they need not maximize the algorithmic firm's profit. The mechanism weakens but does not disappear as the number of non-algorithmic competitors increases. Moreover, competitive outcomes need not be restored by limited strategic sophistication: when rivals adapt prices using simple learning rules based only on their own realized profits, a fast algorithm can still steer behavior toward supracompetitive fixed points using simple pricing rules. High prices do not require rivals to even be aware that they face an algorithmic opponent.

Our framework also has implications for platform design and platform neutrality. When a platform can choose the features of pricing algorithms used by sellers, the optimal design depends on the weights it attaches to consumer surplus and to the profits of different sellers. A platform that places more weight on producer surplus than consumer surplus has an incentive to equip a subset of sellers with fast, partially committed algorithms in order to soften competition without explicit coordination. In this case, constraints on the speed and commitment features of platform-provided pricing tools, or policies that ensure neutral access to such tools, can increase welfare.

Finally, our analysis sheds light on potential policy responses and suggests directions for future work. Restricting firms from rapid and automatic reactions to rivals' prices can mitigate coercion. Even so, persistence of pricing rules alone has the potential to raise prices. In our model, a firm benefits from reacting faster to rivals' prices, creating the possibility of an "arms race" in pricing technology, in which firms invest in ever-faster algorithms with potentially negative welfare effects, much like the high-frequency trading arms race studied by Budish et al. (2015). Remaining challenges for future research include endogenizing the adoption of pricing algorithms and the investment in speed and commitment, studying richer learning and behavioral rules on the part of slower firms, and developing empirical tests that link observed pricing technologies to market outcomes.

References

Abreu, D., P. Milgrom, and D. Pearce (1991). Information and timing in repeated partnerships. *Econometrica*, 1713–1733.

Anufriev, M., D. Kopányi, and J. Tuinstra (2013). Learning cycles in bertrand competition with differentiated commodities and competing learning rules. *Journal of Economic Dynamics and Control 37*(12), 2562–2581.

Aparicio, D., Z. Metzman, and R. Rigobon (2024). The pricing strategies of online grocery retailers. *Quantitative Marketing and Economics 22*(1), 1–21.

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

Assad, S., R. Clark, D. Ershov, and L. Xu (2024). Algorithmic pricing and competition: Empirical evidence from the German retail gasoline market. *Journal of Political Economy 132*(3), 723–771.

Banchio, M. and G. Mantegazza (2022). Artificial intelligence and spontaneous collusion. *arXiv preprint arXiv:2202.05946*.

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

Budish, E., P. Cramton, and J. Shim (2015). The high-frequency trading arms race: Frequent batch auctions as a market design response. *The Quarterly Journal of Economics 130*(4), 1547–1621.

Byrne, D. P., N. de Roos, M. S. Lewis, L. M. Marx, and X. Wu (2025). Asymmetric information sharing in oligopoly: A natural experiment in retail gasoline. *Journal of Political Economy 133*(7).

Calder-Wang, S. and G. H. Kim (2023). Coordinated vs efficient prices: The impact of algorithmic pricing on multifamily rental markets. *Available at SSRN 4403058*.

Calvano, E., G. Calzolari, V. Denicolo, and S. Pastorello (2020). Artificial Intelligence, Algorithmic Pricing, and Collusion. *American Economic Review 110*(10), 3267–97.

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

Federal Trade Commission et al. (2023, November). Complaint [public redacted version with redactions updated]. Federal Trade Commission et al. v. Amazon.com, Inc., Case No. 2:23-cv-01495-JHC (W.D. Wash.), Document 114 (filed Nov. 2, 2023).

FeedbackWhiz (2025). Pros and cons of Amazon automated pricing tool. https://www.feedbackwhiz.com/blog/pros-and-cons-of-amazons-automate-pricing-tool/. Accessed February 2025.

Fudenberg, D., D. M. Kreps, and E. S. Maskin (1990). Repeated Games with Long-Run and Short-Run Players. *The Review of Economic Studies 57*(4), 555.

Fudenberg, D. and D. K. Levine (1989). Reputation and equilibrium selection in games with a patient player. *Econometrica 57*(4), 759–778.

Fudenberg, D. and D. K. Levine (2016). Whither game theory? towards a theory of learning in games. *Journal of Economic Perspectives 30*(4), 151–170.

Fudenberg, D. and J. Tirole (1986). A "signal-jamming" theory of predation. *The RAND Journal of Economics*, 366–376.

Gale, D. and H. Sabourian (2005). Complexity and competition. *Econometrica 73*(3), 739–769.

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

Harrington, J. E. (2022). The effect of outsourcing pricing algorithms on market competition. *Management Science 68*(9), 6889–6906.

Harrington, J. E. (2025). An economic test for an unlawful agreement to adopt a third-party's pricing algorithm. *Economic Policy 40*(121), 261–295.

Informed Software (2025). Informed repricer: Premium repricer for Amazon sellers. https://www.informedrepricer.com/. Accessed February 2025.

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

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

Kreps, D. M. and R. Wilson (1982). Reputation and imperfect information. *Journal of Economic Theory 27*(2), 253–279.

Lamba, R. and S. Zhuk (2025). Pricing with algorithms. *arXiv preprint arXiv:2205.04661*.

Leisten, M. (2024). Algorithmic competition, with humans. Working paper.

Levine, D. K. (2024). Efficiently breaking the folk theorem by reliably communicating long term commitments. *Available at SSRN*. Working paper.

Maskin, E. and J. Tirole (1988). A Theory of Dynamic Oligopoly, I: Overview and Quantity Competition with Large Fixed Costs. *Econometrica 56*(3), 549–569.

Miklós-Thal, J. and C. Tucker (2019). Collusion by Algorithm: Does Better Demand Prediction Facilitate Coordination Between Sellers? *Management Science 65*(4), 1552–1561.

Milgrom, P. and J. Roberts (1982). Predation, reputation, and entry deterrence. *Journal of Economic Theory 27*(2), 280–312.

Musolff, L. (2024). Algorithmic pricing facilitates tacit collusion. Working paper.

O'Connor, J. and N. E. Wilson (2021). Reduced Demand Uncertainty and the Sustainability of Collusion: How AI Could Affect Competition. *Information Economics and Policy 54*.

Porter, R. H. (2005). Detecting collusion. *Review of Industrial Organization 26*(2), 147–167.

Repricer.com (2023). Repricer.com: The fastest Amazon repricer. `https://www.repricer.com/`. Accessed October 2023.

Retail Cloud (2025). Dynamic pricing in retail: How ai is helping small businesses maximize margins. `https://retailcloud.com/dynamic-pricing-in-retail/`. Accessed October 2025.

Salcedo, B. (2015). Pricing Algorithms and Tacit Collusion. Working paper.

Salop, S. (1986). Practices that (Credibly) Facilitate Oligopoly Coordination. In J. Stiglitz and F. Mathewson (Eds.), *New Developments in the Analysis of Market Structure*. MIT Press, Cambridge, MA.

Singh, N. and X. Vives (1984). Price and Quantity Competition in a Differentiated Duopoly. *The RAND Journal of Economics 15*(4), 546–554.

Sugaya, T., G. Stanford, and A. Wolitzky (2025). Collusion with optimal information disclosure. Technical report, Working Paper.

Taiga Data (2025). Competitive fuel pricing. `https://www.taigadata.com/front-office-platform/competitive-fuel-pricing/`. Accessed February 2025.

Tirole, J. (1988). *The Theory of Industrial Organization*. MIT Press, Cambridge, MA.

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

Appendix

A.1 Proofs

Proof of Proposition 1

Before proving the proposition, we formally define the rival's punished deviation payoff and the set of incentive-compatible prices. We introduce some additional notation: Let $P_i \subset \mathbb{R}$ denote the compact feasible set of prices for firm $i \in \{a, b\}$. Write the commitment parameter as $\delta \equiv \beta\gamma \in [0, 1)$ and define the continuation weight as $\omega(\delta) \equiv \delta/(1-\delta)$.

Fix a reprogramming date. Under Assumption A2, if firm b deviates to $\hat{p}_b \neq p_b$ it faces p_a for a $(1-\alpha)$ fraction of the period and the punishment $R_a(\hat{p}_b)$ for the remaining α fraction. Thus its best deviation payoff is

$$D(p_a; \alpha) \equiv \max_{\hat{p}_b \in P_b} \Big\{ (1-\alpha)\pi_b(\hat{p}_b, p_a) + \alpha\, \pi_b\big(\hat{p}_b, R_a(\hat{p}_b)\big) \Big\}.$$

Compliance is optimal iff $\pi_b(p_b, p_a) \geq D(p_a; \alpha)$. Define the incentive-compatible set

$$\mathcal{F}(p_a; \alpha) \equiv \{p_b \in P_b :\ \pi_b(p_b, p_a) \geq D(p_a; \alpha)\}.$$

Denote the maximum feasible value $\bar{p}_b(p_a; \alpha) \equiv \max \mathcal{F}(p_a; \alpha)$. Because P_b is compact and π_b and D are continuous, $\mathcal{F}(p_a; \alpha)$ is closed in P_b and hence compact. Therefore, whenever $\mathcal{F}(p_a; \alpha) \neq \emptyset$, $\bar{p}_b(p_a; \alpha)$ exists.

Firm a's objective (given p_{bt}) is

$$\Psi_a(p_a, p_b; p_{bt}) = (1-\alpha)\pi_a(p_a, p_{bt}) + \alpha\, \pi_a(\sigma(p_{bt}), p_{bt}) + \omega(\delta)\pi_a(p_a, p_b),$$

where $\sigma(p_{bt}) = p_a$ if $p_b = p_{bt}$ and $\sigma(p_{bt}) = R_a(p_{bt})$ otherwise.

We provide sufficient conditions for equilibrium existence and uniqueness.

Assumption R1 (Regularity for existence and uniqueness)**.** *Maintain Assumptions A1 and A2. Assume:*

- *(i) The feasible domain $\{p_a \in P_a : \mathcal{F}(p_a; \alpha) \neq \emptyset\}$ is a nonempty compact interval, and for every $p_{bt} \in P_b$ the reduced-form objective $p_a \mapsto \Psi_a\big(p_a, p_b; p_{bt}\big)$ is continuous and strictly quasi-concave on this domain.*
- *(ii) The mapping $p_{bt} \mapsto p_b^{\dagger}(p_{bt})$ is continuously differentiable and satisfies $\sup_{p_{bt} \in P_b} \left| \frac{\partial p_b^{\dagger}}{\partial p_{bt}} \right| < 1$.*
- *(iii) For any $p_{bt} \in P_b$ and any p_a with $\mathcal{F}(p_a; \alpha) \neq \emptyset$, firm a selects the largest incentive-compatible continuation price for firm b: $p_b = \bar{p}_b(p_a; \alpha)$.*

These conditions admit natural economic interpretations and are satisfied in many standard differentiated-products pricing environments. Condition (i) is a single-peakedness requirement: for any rival price p_{bt}, firm a's (one-dimensional) reduced-form target problem is well behaved, implying a unique optimizer $p_a^\dagger(p_{bt})$ when the induced target $p_b^\dagger(p_{bt})$ is pinned down as a function of $p_a^\dagger(p_{bt})$. Condition (ii) is a contraction requirement that rules out excessive strategic feedback and ensures a unique fixed point. Demand systems that yield strict quasi-concavity of each firm's period profit with respect to own price together with cross-price effects that are small relative to own-price effects satisfy these conditions, as such primitives ensure that best responses and the incentive-compatible boundary $\bar{p}_b(\cdot;\alpha)$ are well-behaved. For example, linear and multinomial logit demand systems satisfy these requirements under standard parameter restrictions, which can be verified via a diagonal-dominance condition.[21]

Condition (iii) ensures stability in firm a's choice of the continuation target for firm b. In general, the continuation term $\omega(\delta)\pi_a(p_a, p_b)$ is strictly increasing in p_b (substitutes), which is the primary economic rationale for a to choose $p_b = \bar{p}_b(p_a;\alpha)$. However, when $p_{bt} = \bar{p}_b(p_a;\alpha)$, firm a may consider a small deviation downward in p_b to be profitable. Because $\sigma(\cdot)$ treats the event $p_b = p_{bt}$ differently from $p_b \neq p_{bt}$, reducing p_b to force firm b to be non-compliant (even when it desires to comply) can discontinuously improve the α-weighted term while reducing the continuation payoff by an arbitrarily small amount. In this case, the reduced-form program may not be well-defined. Condition (iii) rules out this pathology.

This is primarily a convenience assumption to ensure intuitive economic behavior. It is unnecessary in the high-commitment limit $\delta = \beta\gamma \to 1$. The behavior can be microfounded by augmenting the model to use a discrete price grid, a small compliance band (rather than a point trigger), imposing a vanishing cost of triggering the within-period punishment phase, or by using alternative timing assumptions, as we discuss in Appendix A.2.

Proof. For any feasible p_a, firm a selects the maximal incentive-compatible continuation price for b, i.e., $p_b = \bar{p}_b(p_a;\alpha)$, by Assumption R1(iii). Hence firm a's problem reduces to the one-dimensional program

$$\max_{p_a:\,\mathcal{F}(p_a;\alpha)\neq\emptyset} \Psi_a\big(p_a, \bar{p}_b(p_a;\alpha); p_{bt}\big).$$

By Assumption R1(i), the feasible set is an interval and the reduced objective is continuous and strictly quasi-concave, so the maximizer $p_a^\dagger(p_{bt})$ is unique. Then $p_b^\dagger(p_{bt}) = \bar{p}_b(p_a^\dagger(p_{bt});\alpha)$ is unique as well. This establishes a unique target pair conditional on any p_{bt}.

In equilibrium, the current price must coincide with the programmed target, so p_{bt} satisfies the fixed-point condition

$$p_{bt} = p_b^\dagger(p_{bt}).$$

[21]Demand systems with flexible curvature, such as random-coefficients (mixed) logit, do not necessarily have globally quasi-concave profit functions. In those settings, one typically imposes additional curvature restrictions (e.g., that the distribution of types is log-concave) or verifies regularity numerically.

By Assumption R1(ii), the mapping $p_{bt} \mapsto p_b^\dagger(p_{bt})$ is continuously differentiable with slope strictly less than one in absolute value. Hence it is a contraction on P_b, and the Banach fixed-point theorem implies that there exists a unique $p_{bt}^* \in P_b$ satisfying $p_{bt}^* = p_b^\dagger(p_{bt}^*)$. Thus an equilibrium price p_{bt}^* exists and is unique.

Following the result from the first part of the proof, there exists a unique optimal target pair $(p_a^*, p_b^*) = (p_a^\dagger(p_{bt}^*), p_b^\dagger(p_{bt}^*))$. Therefore, the equilibrium (p_{bt}^*, p_a^*, p_b^*) exists and is unique. □

Proof of Proposition 2

Proof. Set $\alpha = 0$. Then the algorithm has no within-period update, so firm a charges $p_a = \rho$ throughout the period. Firm b's static problem in Assumption A1 reduces to $\max_{p_b} \pi_b(p_b, \rho)$, implying the unique best response $p_b = R_b(\rho)$. Hence any equilibrium price vector satisfying Assumption A2 must lie on firm b's best-response function: $(p_a^\dagger, p_b^\dagger) = (\rho^\dagger, R_b(\rho^\dagger))$.

At a reprogramming date, firm a chooses ρ taking the current rival price p_{bt} as given (as in $V_0(t)$), but it internalizes that if the algorithm remains in force in future periods, firm b will best respond to ρ. The effective discount factor for remaining committed is $\beta\gamma$, so the discounted weight on future committed periods is $\sum_{k\geq 1}(\beta\gamma)^k = \beta\gamma/(1-\beta\gamma)$. Therefore $\rho^\dagger$ solves

$$\max_{\rho \in P_a} \left\{ \pi_a(\rho, p_{bt}) + \frac{\beta\gamma}{1-\beta\gamma} \pi_a(\rho, R_b(\rho)) \right\}. \tag{A.1}$$

In equilibrium, $p_{bt} = p_b^\dagger = R_b(\rho^\dagger)$.

Let ρ^B denote firm a's Bertrand–Nash price (simultaneous price setting), characterized by $\partial_1 \pi_a(\rho^B, R_b(\rho^B)) = 0$ (equivalently $\rho^B = R_a(R_b(\rho^B))$), and let ρ^L denote firm a's Stackelberg-leader price, characterized by

$$\frac{d}{d\rho}\pi_a(\rho, R_b(\rho))\Big|_{\rho=\rho^L} = \partial_1 \pi_a(\rho^L, R_b(\rho^L)) + \partial_2 \pi_a(\rho^L, R_b(\rho^L)) R_b'(\rho^L) = 0. \tag{A.2}$$

An interior optimum $\rho^\dagger$ of the objective satisfies the first-order condition

$$0 = \partial_1 \pi_a(\rho^\dagger, p_b^\dagger) + \frac{\beta\gamma}{1-\beta\gamma}\Big[\partial_1 \pi_a(\rho^\dagger, R_b(\rho^\dagger)) + \partial_2 \pi_a(\rho^\dagger, R_b(\rho^\dagger)) R_b'(\rho^\dagger)\Big].$$

Substituting $p_b^\dagger = R_b(\rho^\dagger)$ and simplifying yields

$$\partial_1 \pi_a(\rho^\dagger, R_b(\rho^\dagger)) + \beta\gamma\, \partial_2 \pi_a(\rho^\dagger, R_b(\rho^\dagger)) R_b'(\rho^\dagger) = 0.$$

Define $F(\rho) \equiv \partial_1 \pi_a(\rho, R_b(\rho)) + \beta\gamma\, \partial_2 \pi_a(\rho, R_b(\rho)) R_b'(\rho)$. Then $F(\rho^B) = \beta\gamma\, \partial_2 \pi_a(\rho^B, R_b(\rho^B)) R_b'(\rho^B) > 0$, while $F(\rho^L) = (\beta\gamma - 1)\partial_2 \pi_a(\rho^L, R_b(\rho^L)) R_b'(\rho^L) < 0$ since $\beta\gamma \in (0,1)$ and products are substitutes and prices are strategic complements ($\partial_2 \pi_a > 0$ and $R_b' > 0$). By continuity, there exists $\rho^\dagger \in (\rho^B, \rho^L)$ solving $F(\rho^\dagger) = 0$. Under the regularity conditions ensuring strict quasi-concavity of the objective, this solution is unique. Since R_b is increasing, the induced $p_b^\dagger = R_b(\rho^\dagger)$ lies between $R_b(\rho^B)$ and $R_b(\rho^L)$, completing the proof. $\square$

Proof of Proposition 3

Proof. Fix an arbitrary commitment level $\delta \equiv \beta\gamma \in [0,1)$ and consider two speed levels $\alpha' > \alpha$. Let $(p_a^\dagger, p_b^\dagger)$ denote the coercive equilibrium target prices under (α, δ).

First, we establish that the rival's deviation payoff decreases in α. For any deviation price $\hat{p}_b \in P_b \setminus p_b^\dagger$ and any on-path $p_a^\dagger$, define

$$d(\hat{p}_b; p_a^\dagger, \alpha) \equiv (1-\alpha)\pi_b(\hat{p}_b, p_a^\dagger) + \alpha\, \pi_b\big(\hat{p}_b, R_a(\hat{p}_b)\big),$$

so that the punished deviation payoff is $D(p_a^\dagger; \alpha) = \sup_{\hat{p}_b} d(\hat{p}_b; p_a^\dagger, \alpha)$. Then

$$d(\hat{p}_b; p_a^\dagger, \alpha') - d(\hat{p}_b; p_a^\dagger, \alpha) = (\alpha' - \alpha)\Big[\pi_b\big(\hat{p}_b, R_a(\hat{p}_b)\big) - \pi_b(\hat{p}_b, p_a^\dagger)\Big].$$

Under Assumption A2, $\sigma(\hat{p}_b) = R_a(\hat{p}_b)$ for $\hat{p}_b \neq p_b^\dagger$. By the hypothesis of the proposition $\pi_b(\hat{p}_b, R_a(\hat{p}_b)) = \pi_b(\hat{p}_b, \sigma(\hat{p}_b)) \leq \pi_b(\hat{p}_b, p_a^\dagger)$ for all deviations $\hat{p}_b \neq p_b^\dagger$. Hence $d(\hat{p}_b; p_a^\dagger, \alpha') \leq d(\hat{p}_b; p_a^\dagger, \alpha)$ for all $\hat{p}_b$, and therefore

$$D(p_a^\dagger; \alpha') \leq D(p_a^\dagger; \alpha).$$

Next, we show that any equilibrium target pair feasible at α remains feasible at α'. Because $(p_a^\dagger, p_b^\dagger)$ is an equilibrium target pair at speed α, it satisfies firm b's incentive constraint

$$\pi_b(p_b^\dagger, p_a^\dagger) \geq D(p_a^\dagger; \alpha).$$

Using our earlier result that $D(p_a^\dagger; \alpha') \leq D(p_a^\dagger; \alpha)$ yields

$$\pi_b(p_b^\dagger, p_a^\dagger) \geq D(p_a^\dagger; \alpha'),$$

so the same target pair is incentive compatible for firm b at speed α'.

Finally, we establish that firm a can secure at least its α-equilibrium profit when speed is α'. Consider the environment with speed α' and the same commitment level δ. Firm a can choose the same target-price algorithm that implements $(p_a^\dagger, p_b^\dagger)$ on path and uses $R_a(\cdot)$ off path. Because we have shown that the target pair is feasible and compatible with firm b's incentive compatibility constraint, firm b complies and chooses $p_b^\dagger$, the update is not triggered, and firm a charges $p_a^\dagger$ throughout the period, yielding the same flow profit $\pi_a(p_a^\dagger, p_b^\dagger)$. Because equal profits are feasible for $\alpha' > \alpha$, equilibrium profit is weakly increasing in the speed advantage α for any $\delta \in [0,1)$. $\square$

Proof of Proposition 4

Proof. Fix α and write $\delta \equiv \beta\gamma \in [0,1)$. At a reprogramming date, conditional on the current rival price p_{bt}, firm a chooses a target-price algorithm (equivalently, a feasible target pair $(p_a^\dagger, p_b^\dagger)$ together with the on/off-path implementation constraints) to maximize

$$(1-\alpha)\pi_a(\rho, p_{bt}) + \alpha\pi_a(\sigma(p_{bt}), p_{bt}) + \frac{\delta}{1-\delta}\,\pi_a(p_a^\dagger, p_b^\dagger), \tag{A.3}$$

subject to the implementation constraints (i)–(ii) and firm b's incentive constraint (iii). Importantly, the feasible set does not depend on δ.

Take $\delta' > \delta$. For any fixed feasible choice, the first two (current-period) terms do not depend on δ, while $\delta \mapsto \frac{\delta}{1-\delta}$ is strictly increasing on $[0,1)$. Hence the objective is weakly increasing in δ for every feasible choice whenever $\pi_a(p_a^\dagger, p_b^\dagger) \geq 0$. Therefore the maximized value (firm a's equilibrium discounted payoff) is weakly increasing in δ and is strictly increasing whenever the equilibrium on-path flow profit is strictly positive.

Moreover, the firm's objective can be interpreted as maximizing the weighted sum of current payoffs $(1-\alpha)\pi_a(\rho, p_{bt}) + \alpha\pi_a(\sigma(p_{bt}), p_{bt})$ and future payoffs $\pi_a(p_a^\dagger, p_b^\dagger)$. The term $\pi_a(p_a^\dagger, p_b^\dagger)$ yields the on-path equilibrium profit, and the weight on this component is increasing with δ. Since the set of feasible target prices does not depend on δ, then, by revealed preference, the component $\pi_a(p_a^\dagger, p_b^\dagger)$ is weakly increasing with δ; otherwise, a preferred solution would have been available, contradicting optimality. Therefore, the per-period profits are weakly increasing in δ. □

Proof of Proposition 5

Proof. Denote the joint profit-maximizing (collusive) prices (p_a^C, p_b^C) and the sequential-move prices where firm b is the leader as (p_a^S, p_b^S), i.e.,

$$p_b^S = \arg\max_{p_b} \pi_b(p_b, R_a(p_b)) \tag{A.4}$$

$$p_a^S = R_a(p_b^S) \tag{A.5}$$

and $R_a(p_b)$ provides firm a's static best-response function. The collusive prices satisfy $(p_a^C, p_b^C) = \arg\max_{(p_a,p_b)} [\pi_a(p_a, p_b) + \pi_b(p_b, p_a)]$ and yield the first-order conditions $\frac{\partial \pi_a}{\partial p_a} + \frac{\partial \pi_b}{\partial p_a} = 0$, $\frac{\partial \pi_b}{\partial p_b} + \frac{\partial \pi_a}{\partial p_b} = 0$.

It is sufficient to consider $\alpha = 1$ and the full-commitment limit $\delta \equiv \beta\gamma \to 1$. In that limit, firm a's choice of the target pair solves

$$\max_{(p_a^\dagger, p_b^\dagger)} \; \pi_a(p_a^\dagger, p_b^\dagger) \tag{A.6}$$

$$\text{s.t. } \pi_b(p_b^\dagger, p_a^\dagger) \geq \max_{p_b} \pi_b(p_b, R_a(p_b)). \tag{A.7}$$

Assume $\pi_b(p_b^C, p_a^C) > \max_{p_b} \pi_b(p_b, R_a(p_b)) = \pi_b(p_b^S, p_a^S)$. Then $(p_a^\dagger, p_b^\dagger) = (p_a^C, p_b^C)$ is feasible and the constraint is slack. By continuity of π_b, there exists $\varepsilon > 0$ such that $(p_a^C - \varepsilon, p_b^C)$ remains feasible. Moreover, at (p_a^C, p_b^C) we have $\partial \pi_a / \partial p_a < 0$ (from the collusive FOC and $\partial \pi_b / \partial p_a > 0$), so for $\varepsilon > 0$ small, $\pi_a(p_a^C - \varepsilon, p_b^C) > \pi_a(p_a^C, p_b^C)$. Hence firm a can achieve profit strictly above its profit under collusive prices in the limiting problem.

Finally, by the maximum theorem and under the maintained uniqueness/regularity conditions, equilibrium target prices and profits vary continuously in (α, δ) in a neighborhood of $(1,1)$. Therefore the strict inequality persists for all (α, δ) sufficiently close to $(1,1)$, implying the existence of thresholds $\overline{\alpha} < 1$ and $\overline{\delta} < 1$ such that for all $\alpha > \overline{\alpha}$ and $\beta\gamma > \overline{\delta}$, firm a's equilibrium profit is strictly greater than $\pi_a(p_a^C, p_b^C)$. □

Proof of Proposition 6

Proof. Fix rival price p_{bt}. Write the commitment parameter as $\delta \equiv \beta\gamma \in [0,1)$ and the continuation weight as $\omega(\delta) \equiv \delta/(1-\delta)$.

For any candidate target price p_a, define firm b's maximal deviation payoff

$$W(\alpha; p_a) \;\equiv\; \max_{\hat{p}_b} \Big\{ (1-\alpha)\pi_b(\hat{p}_b, p_a) \;+\; \alpha\, \pi_b\big(\hat{p}_b, R_a(\hat{p}_b)\big) \Big\}, \tag{A.8}$$

and let $\hat{p}_b(\alpha, p_a)$ denote the (unique) maximizer. Under linear demand, $\hat{p}_b(\alpha, p_a)$ is unique and $W(\alpha; p_a)$ is continuously differentiable. By the envelope theorem,

$$\frac{\partial W(\alpha; p_a)}{\partial \alpha} = \pi_b\big(\hat{p}_b, R_a(\hat{p}_b)\big) - \pi_b(\hat{p}_b, p_a) \;<\; 0, \tag{A.9}$$

where the strict inequality follows from the fact that the binding deviation is downward and the best-response punishment is sufficiently severe.

In the coercive solution, firm a's objective is strictly increasing in the target $p_b^\dagger$ (products are substitutes), so the incentive compatibility constraint binds:

$$\pi_b(p_b^\dagger, p_a^\dagger) = W(\alpha; p_a^\dagger). \tag{A.10}$$

Define $F(p_a, p_b, \alpha) \equiv \pi_b(p_b, p_a) - W(\alpha; p_a)$. At the coercive solution, $p_b^\dagger$ lies on the decreasing side of firm b's single-peaked profit in own price, so $F_{p_b}(p_a^\dagger, p_b^\dagger, \alpha) = \pi_{b,1}(p_b^\dagger, p_a^\dagger) < 0$. Hence, by the implicit function theorem there exists a differentiable *IC frontier* $p_b = \varphi(p_a, \alpha)$ in a neighborhood of $(p_a^\dagger, \alpha)$ satisfying $F(p_a, \varphi(p_a, \alpha), \alpha) = 0$.

Moreover, $\varphi_\alpha(p_a, \alpha) > 0$: increasing α strictly lowers the deviation value $W(\alpha; p_a)$, so maintaining the binding equality requires a (weakly) higher p_b along the IC frontier. Under linear demand, the same implicit-function argument implies $\varphi_{p_a}(p_a, \alpha) > 0$ as well: increasing p_a raises on-path profit for firm b more than it raises deviation profit (the punishment portion of deviation payoff does not depend on p_a), so the binding IC frontier shifts upward in p_a.

Using that π_a is increasing in the rival's price, firm a chooses the *largest* incentive-compatible p_b for any p_a, so $p_b^\dagger = \varphi(p_a^\dagger, \alpha)$. Substituting into the target-price problem *(7)* yields the reduced problem

$$\max_{p_a} \left\{ \Pi^{cur}(p_a; \alpha, p_{bt}) \;+\; \omega(\delta)\, \pi_a\big(p_a, \varphi(p_a, \alpha)\big) \right\}, \tag{A.11}$$

where $\Pi^{cur}(p_a; \alpha, p_{bt}) \equiv (1-\alpha)\pi_a(p_a, p_{bt}) + \alpha\, \pi_a(\sigma(p_{bt}), p_{bt})$ collects the current-period term in *(7)* (which does not depend on δ). Under linear demand, the objective is strictly concave in p_a, so the maximizer $p_a^\dagger(\alpha, \delta)$ is unique and satisfies a unique first-order condition $G(p_a, \alpha, \delta) = 0$ with $G_{p_a} < 0$. Therefore, $p_a^\dagger(\alpha, \delta)$ is differentiable in (α, δ) by the implicit function theorem.

To sign comparative statics, note first that at the optimum we must have $\partial \Pi^{cur}/\partial p_a \leq 0$: if instead $\partial \Pi^{cur}/\partial p_a > 0$, then a small increase in p_a would strictly raise the current-period term and (because $\varphi_{p_a} > 0$ and π_a is increasing in p_b) would also weakly raise the continuation term, contradicting optimality. Since the first-order condition sets

$$\frac{\partial \Pi^{cur}}{\partial p_a} + \omega(\delta)\, \frac{\partial}{\partial p_a} \pi_a\big(p_a, \varphi(p_a, \alpha)\big) = 0, \tag{A.12}$$

it follows that the marginal continuation payoff $\frac{\partial}{\partial p_a} \pi_a(p_a, \varphi(p_a, \alpha))$ is weakly nonnegative at $p_a^\dagger$.

Now consider increases in the parameters: (i) increasing δ increases $\omega(\delta)$, placing more weight on a continuation term with weakly nonnegative marginal payoff, so the unique optimizer $p_a^\dagger$ is weakly increasing in δ; (ii) increasing α shifts the IC frontier upward, and because π_a exhibits strategic complements under linear demand, this weakly raises the marginal continuation benefit of increasing p_a. In addition, increasing α weakly reduces the weight on the portion of Π^{cur} that depends on p_a, which (since $\partial \Pi^{cur}/\partial p_a \leq 0$ at the optimum) also shifts the first-order condition in a direction that weakly increases the optimizer. Hence $p_a^\dagger$ is weakly increasing in α as well.

Finally, since $p_b^\dagger(\alpha, \delta) = \varphi(p_a^\dagger(\alpha, \delta), \alpha)$ with $\varphi_{p_a} > 0$ and $\varphi_\alpha > 0$, the target $p_b^\dagger$ inherits the same monotonicity: it is weakly increasing in α and (through $p_a^\dagger$) weakly increasing in $\delta = \beta\gamma$. $\square$

Proof of Proposition 7

Proof. Let $CS(p_a, p_b)$ denote consumer surplus. By Roy's identity, $\partial CS/\partial p_i = -D_i(p_a, p_b) < 0$ for $i \in \{a, b\}$. By Proposition 6, both equilibrium target prices $p_a^\dagger$ and $p_b^\dagger$ are weakly increasing in α and in $\beta\gamma$. Therefore, consumer surplus is weakly decreasing in α and in $\beta\gamma$. $\square$

Proof of Proposition 8

Proof. Maintain the linear-demand specification (22) with marginal costs normalized to zero and the restriction to nonnegative prices $p_b \geq 0$ used in Section 5.2. Under the restriction that firm a commits to a linear rule through the origin and the target point $(p_b^\dagger, p_a^\dagger)$, the within-period update can be written as

$$\sigma(p_b) = \phi\, p_b, \qquad \text{where} \qquad \phi = \frac{p_a^\dagger}{p_b^\dagger} = \frac{(2d+1)(3d+1)}{(\alpha+5)d(d+1)+1} > 0. \tag{A.13}$$

Since $\sigma(0) = 0$, the nonnegativity truncation in (25) never binds for $p_b \geq 0$.

For any $p_b \geq 0$, firm b's realized per-period profit (averaging over the $(1-\alpha)$ and α portions of the period) is

$$\tilde{\pi}_b(p_b) \equiv (1-\alpha)\pi_b(p_b, p_a^\dagger) + \alpha\, \pi_b\big(p_b, \sigma(p_b)\big). \tag{A.14}$$

Under (22),

$$\pi_b(p_b, p_a) = p_b\, D_b(p_b, p_a) = p_b\Big[1 - \Big(\tfrac{1}{4} + \tfrac{d}{2}\Big)p_b + \tfrac{d}{2}\, p_a\Big]. \tag{A.15}$$

Substituting $\sigma(p_b) = \phi p_b$ gives

$$\tilde{\pi}_b(p_b) = \underbrace{\Big[1 + \tfrac{d}{2}(1-\alpha)p_a^\dagger\Big]}_{\equiv A}\, p_b - \underbrace{\Big[\tfrac{1}{4} + \tfrac{d}{2} - \tfrac{\alpha d}{2}\phi\Big]}_{\equiv \kappa}\, p_b^2. \tag{A.16}$$

Hence $\tilde{\pi}_b''(p_b) = -2\kappa$ for all p_b. Moreover, using (A.13),

$$\kappa = \frac{(2d+1)\big[(5d^2+5d+1) - \alpha(5d^2+d)\big]}{4\big[(\alpha+5)d(d+1)+1\big]} \geq \frac{(2d+1)(4d+1)}{4\big[(\alpha+5)d(d+1)+1\big]} > 0, \tag{A.17}$$

where the inequality uses $\alpha \in [0,1]$ and $d > 0$. Therefore, $\tilde{\pi}_b$ is strictly concave and single-peaked on $[0, \infty)$. Because $\tilde{\pi}_b$ is strictly concave, it has a unique maximizer

$$p_b^\dagger = \arg\max_{p_b \geq 0} \tilde{\pi}_b(p_b) = \frac{A}{2\kappa}. \tag{A.18}$$

Since $\tilde{\pi}_b'(p_b) = A - 2\kappa p_b$ is affine with strictly negative slope, it follows that

$$\tilde{\pi}_b'(p_b) > 0 \text{ for } p_b < p_b^\dagger, \qquad \tilde{\pi}_b'(p_b) < 0 \text{ for } p_b > p_b^\dagger. \tag{A.19}$$

We next establish convergence. Consider first the idealized update rule (24) applied to $\tilde{\pi}_b$ when the derivative is known: $p_{b,t+1} = p_{bt} + \lambda\, \tilde{\pi}_b'(p_{bt})$. Let $e_t \equiv p_{bt} - p_b^\dagger$. Using $\tilde{\pi}_b'(p) = A - 2\kappa p$ and $p_b^\dagger = A/(2\kappa)$ yields

$$e_{t+1} = (1 - 2\lambda\kappa)\, e_t, \tag{A.20}$$

so $p_{bt} \to p_b^\dagger$ for any initial p_{b0} whenever $0 < \lambda\kappa < 1$ (equivalently $0 < \lambda < 1/\kappa$).

Now consider the finite-difference estimator (29) used in the text, with $p_b^0 \neq p_b^1$ so the secant slope is well-defined. In the deterministic environment of the proposition, realized profits satisfy $\pi_{bt}^* = \tilde{\pi}_b(p_{bt})$. Since $\tilde{\pi}_b$ is quadratic, the secant slope equals the derivative at the midpoint:

$$\frac{\tilde{\pi}_b(p_{bt}) - \tilde{\pi}_b(p_{b,t-1})}{p_{bt} - p_{b,t-1}} = \tilde{\pi}_b'\!\left(\frac{p_{bt} + p_{b,t-1}}{2}\right). \tag{A.21}$$

Substituting this into (24) implies the two-dimensional linear recursion

$$e_{t+1} = (1 - \lambda\kappa)e_t - \lambda\kappa\, e_{t-1}. \tag{A.22}$$

Its characteristic polynomial is $r^2 - (1 - \lambda\kappa)r + \lambda\kappa = 0$, whose roots lie strictly inside the unit circle whenever $0 < \lambda\kappa < 1$. Hence $e_t \to 0$ (and thus $p_{bt} \to p_b^\dagger$) for any initial pair (p_b^0, p_b^1) with $p_b^0 \neq p_b^1$ and convergence is attained, provided λ is not too large. □

Proof of Proposition 9

Proof. When $\kappa_{CS} = \kappa_a = \kappa_b \equiv \kappa > 0$, the platform objective (31) is proportional to

$$CS^*(\alpha, \gamma) + \pi_a^*(\alpha, \gamma) + \pi_b^*(\alpha, \gamma), \tag{A.23}$$

i.e., per-period total surplus.

Under the linear demand system (22) (derived from a quasilinear quadratic utility) and with marginal costs normalized to zero, total surplus at a price pair (p_a, p_b) can be written as

$$TS(p_a, p_b) \;=\; CS(p_a, p_b) + \pi_a(p_a, p_b) + \pi_b(p_a, p_b) \;=\; 4 - \frac{1}{8}\big(p_a^2 + p_b^2\big) - \frac{d}{4}\big(p_a - p_b\big)^2. \tag{A.24}$$

Hence $TS(p_a, p_b)$ is strictly decreasing in both $p_a^2 + p_b^2$ and $(p_a - p_b)^2$.

Let $(p_a^*(\alpha, \gamma), p_b^*(\alpha, \gamma))$ denote the coercive equilibrium target prices induced by (α, γ). Proposition 6 (linear-demand comparative statics) implies that both equilibrium prices are weakly increasing in α and in effective commitment γ. In particular, the smallest equilibrium prices occur at $(\alpha, \gamma) = (0, 0)$, which coincides with the simultaneous Bertrand benchmark in the symmetric linear-demand example.

Therefore,

$$TS^*(\alpha, \gamma) = TS(p_a^*(\alpha, \gamma), p_b^*(\alpha, \gamma)) \;\leq\; TS(p_a^*(0, 0), p_b^*(0, 0)) = TS^*(0, 0), \tag{A.25}$$

so $(\alpha,\gamma)=(0,0)$ maximizes the platform objective. $\square$

Proof of Proposition 10

Proof. Let the platform payoff be

$$\Omega(\alpha,\gamma) \equiv \kappa_{CS} CS^*(\alpha,\gamma) + \kappa_a \pi_a^*(\alpha,\gamma) + \kappa_b \pi_b^*(\alpha,\gamma). \tag{A.26}$$

Consider the difference between maximal technology and the no-technology benchmark:

$$\Delta(\kappa_a) \equiv \Omega(1,1) - \Omega(0,0). \tag{A.27}$$

Expanding,

$$\Delta(\kappa_a) = \kappa_{CS}\big(CS^*(1,1) - CS^*(0,0)\big) + \kappa_a\big(\pi_a^*(1,1) - \pi_a^*(0,0)\big) + \kappa_b\big(\pi_b^*(1,1) - \pi_b^*(0,0)\big). \tag{A.28}$$

By Proposition 6, $CS^*(1,1) \leq CS^*(0,0)$. By Propositions 3–4, $\pi_a^*(\alpha,\gamma)$ is weakly increasing in both α and γ. To see that strict monotonicity holds, note that $\pi_a^*(0,0)$ corresponds to the Bertrand profit. By Proposition 5, $\pi_a^*(1,1)$ strictly exceeds the symmetric collusive profit, which is strictly higher than the Bertrand profit under linear demand. Thus, $\pi_a^*(1,1) - \pi_a^*(0,0) > 0$. Hence $\Delta(\kappa_a)$ is affine in κ_a with strictly positive slope.

Define the threshold

$$\bar{\kappa}_a(\kappa_{CS},\kappa_b) \equiv \max\left\{ \frac{\kappa_{CS}\big(CS^*(0,0) - CS^*(1,1)\big) + \kappa_b\big(\pi_b^*(0,0) - \pi_b^*(1,1)\big)}{\pi_a^*(1,1) - \pi_a^*(0,0)}, 0 \right\}. \tag{A.29}$$

Then for any $\kappa_a > \bar{\kappa}_a(\kappa_{CS},\kappa_b)$, we have $\Delta(\kappa_a) > 0$, so the platform strictly prefers $(1,1)$ to $(0,0)$. Therefore $(0,0)$ cannot be optimal and any optimum satisfies $(\alpha^*,\gamma^*) \neq (0,0)$.

Finally, as κ_a becomes large relative to (κ_{CS},κ_b), the objective is dominated by $\kappa_a \pi_a^*(\alpha,\gamma)$. Since $\pi_a^*(\alpha,\gamma)$ is increasing in both α and γ (Propositions 4–3), the platform chooses $(\alpha^*,\gamma^*) = (1,1)$ in the limit. $\square$

Proof of Proposition 11

Proof. Let the platform payoff be

$$\Omega(\alpha,\gamma) \equiv \kappa_{CS} CS^*(\alpha,\gamma) + \kappa_a \pi_a^*(\alpha,\gamma) + \kappa_b \pi_b^*(\alpha,\gamma). \tag{A.30}$$

Consider the commitment-only policy $(\alpha,\gamma) = (0,1)$, which corresponds to the sequential benchmark where seller a leads, and compare it to the simultaneous benchmark $(0,0)$.

Under (22), the simultaneous (Bertrand) price is $p^B = 2/(1+d)$ and each firm earns

$$\pi^B(d) = \frac{1+2d}{(1+d)^2}. \tag{A.31}$$

Under $(\alpha, \gamma) = (0, 1)$ (seller a as leader), prices are

$$p_a^S = \frac{6d+2}{2d^2+4d+1}, \qquad p_b^S = \frac{d\,p_a^S + 2}{2d+1}, \tag{A.32}$$

and equilibrium profits are

$$\pi_a^S(d) = \frac{9d^2+6d+1}{4d^3+10d^2+6d+1}, \qquad \pi_b^S(d) = \frac{25d^4+50d^3+35d^2+10d+1}{8d^5+36d^4+56d^3+36d^2+10d+1}. \tag{A.33}$$

A direct comparison yields

$$\pi_a^S(d) - \pi^B(d) = \frac{d^4}{(d+1)^2(2d+1)(2d^2+4d+1)} \ > \ 0, \tag{A.34}$$

and

$$\pi_b^S(d) - \pi^B(d) = \frac{d^3(9d^3+20d^2+12d+2)}{(d+1)^2(2d+1)(2d^2+4d+1)^2} \ > \ 0 \tag{A.35}$$

for all $d > 0$. Hence the weighted producer term strictly increases:

$$\Delta P \equiv \kappa_a(\pi_a^*(0,1) - \pi_a^*(0,0)) + \kappa_b(\pi_b^*(0,1) - \pi_b^*(0,0)) \ > \ 0, \tag{A.36}$$

since $\kappa_a + \kappa_b > 0$.

By Proposition 7, under linear demand consumer surplus is weakly decreasing in commitment and in speed. Since $(0,1)$ yields strictly higher equilibrium prices than $(0,0)$ when $d > 0$, we have $CS^*(0,1) < CS^*(0,0)$ and thus

$$\Delta CS \equiv CS^*(0,0) - CS^*(0,1) \ > \ 0. \tag{A.37}$$

The platform payoff difference between $(0,1)$ and $(0,0)$ is

$$\Omega(0,1) - \Omega(0,0) = \Delta P - \kappa_{CS}\Delta CS. \tag{A.38}$$

Define

$$\bar{\kappa}_{CS}(\kappa_a, \kappa_b) \equiv \frac{\Delta P}{\Delta CS} \ > \ 0. \tag{A.39}$$

Then for any $\kappa_{CS} \in [0, \bar{\kappa}_{CS})$, we have $\Omega(0,1) > \Omega(0,0)$, so $(0,0)$ cannot be optimal. Therefore any optimal platform policy must satisfy $(\alpha^*, \gamma^*) \neq (0,0)$. $\square$

A.2 Alternative Timing Assumptions

This appendix examines the sensitivity of the coercion mechanism to the move order at reprogramming dates. In the baseline model (Section 2), when the algorithm is reprogrammed in period τ, firm a chooses $\mathcal{A}_\tau = (\rho_\tau, \sigma_\tau)$ and firm b chooses $p_{b\tau}$ simultaneously. As a result, firm b cannot condition on the newly chosen algorithm within period τ and reacts only in subsequent periods. We consider four alternative timing structures for the reprogramming period. Throughout, we maintain the baseline primitives and notation: $\pi_i(p_i, p_j)$ denotes static profit, $\alpha \in [0,1]$ is the within-period speed advantage, $\beta \in [0,1)$ is the discount factor, and $\gamma \in [0,1]$ governs algorithm persistence.

We restrict attention to the same class of "trigger" algorithms used in the baseline analysis: firm a chooses a target price for its rival $p_b^\dagger$ and the algorithm initialization price ρ. The algorithm updates prices according to the rule

$$\sigma(p_b) = \begin{cases} \rho & \text{if } p_b = p_b^\dagger, \\ R_a(p_b) & \text{if } p_b \neq p_b^\dagger, \end{cases} \tag{A.40}$$

where $R_a(\cdot)$ is firm a's static best response (Assumption A2).

Model (a): One-Period Implementation Delay. As in our baseline model, we assume that firm a chooses $\mathcal{A}_\tau$ and firm b chooses $p_{b\tau}$ simultaneously at the beginning of a reprogramming date $t = \tau$. However, we assume that the target price response cannot be implemented during reprogramming periods, but only during subsequent periods which the algorithm remains in place. That is, firm a uses its static best-response during the entirety of the reprogramming period τ. Even if firm b were to choose $p_b = p_b^\dagger$ in period τ, firm a does not accommodate by setting $p_a^\dagger$. In periods $t > \tau$ where the algorithm persists, the algorithm operates exactly as in our baseline model.

This may represent a scenario where the firm is "reprogramming" throughout the period, so that it cannot commit not to manually change the price (and thus it follows its static best-response). Or it may represent a scenario where an algorithm might need to first establish a credible threat or "reset" the market by demonstrating the punishment phase before it can successfully induce compliance.

Given an initialization price ρ_τ for the first $(1-\alpha)$ fraction of the reset period and the forced punishment mapping in the final α fraction, firm b's reprogramming-period best response solves

$$p_{b\tau}^* \in \arg\max_{p_b} \Big[(1-\alpha)\,\pi_b(p_b, \rho_\tau) + \alpha\,\pi_b\big(p_b, R_a(p_b)\big)\Big]. \tag{A.41}$$

In periods $t > \tau$ when the algorithm persists, firm b can comply by choosing $p_b = p_b^\dagger$, which yields $p_a = \rho_\tau$ throughout the period. Denoting this as the target price $p_a^\dagger$, we see that the continuation target is pinned down by the same within-period incentive constraint as in the baseline:

$$\pi_b(p_b^\dagger, p_a^\dagger) \geq \max_{\hat{p}_b} \Big[(1-\alpha)\pi_b(\hat{p}_b, p_a^\dagger) + \alpha\pi_b\big(\hat{p}_b, R_a(\hat{p}_b)\big)\Big]. \tag{A.42}$$

Firm a's objective function differs from the baseline in that it will always follow R_a during the update phase in the reprogramming period and that it has to account for firm b's adjusted IC constraint in this period. Unlike our baseline model, the equilibrium prices may change from the reprogramming period to subsequent periods when the algorithm is in force.

Model (b): No Punishment in the Reprogramming Period. As in our baseline model, firm a chooses $\mathcal{A}_\tau$ and firm b chooses $p_{b\tau}$ simultaneously at the beginning of a reprogramming date $t = \tau$. Here, assume that the within-period update technology $\sigma(\cdot)$ is not operative during the reprogramming period, but instead activates only for future periods $t > \tau$ when the algorithm persists. Thus, firm a's choice of ρ_τ determines its price for the entire period τ. This represents situations where an algorithm needs time to gather data on a rival after it is reprogrammed, or where technical latency prevents immediate response. This "no punishment" in the initial period assumption is the converse to the "always punish" in the initial period assumption in Model (a).

Maintaining the convention that firm a attempts to implement the target immediately, $\rho = p_a^\dagger$, firm b's reprogramming-period best response satisfies

$$p_{bt}^* \in \arg\max_{p_b} \pi_b(p_b, p_a^\dagger) = R_b(p_a^\dagger), \tag{A.43}$$

and the continuation target again must satisfy (A.42) when the update rule becomes active in subsequent periods. Because of the changes in incentives, the equilibrium prices may change from the reprogramming period to the future period when the update rule is active.

Model (c): Target Price Set First, then Simultaneous Price Setting. We break from our baseline model by introducing sequential moves at the beginning of the period. Here, assume that firm a can first commit to the target price $p_b^\dagger$, then (immediately after) firm a chooses ρ and firm b chooses p_b simultaneously. This explores a situation where the compliance price is stickier than firm a's choice of its own price, and may map to scenarios where firm a has to "communicate" its desired outcome to firm b before both firms set price.

In this formulation, the continuation target is pinned down by the incentive constraint (A.42) above. Given a target, the within-period outcome (ρ^*, p_{bt}^*) in the reprogramming period is determined by the Nash equilibrium of the pricing subgame:

$$\rho^* \in \arg\max_{\rho|p_b^\dagger} \Big[(1-\alpha)\pi_a(\rho, p_{bt}^*) + \alpha\pi_a\big(\sigma(p_{bt}^*), p_{bt}^*\big) + \frac{\beta\gamma}{1-\beta\gamma}\,\pi_a(\rho, p_b^\dagger)\Big], \tag{A.44}$$

$$p_{bt}^* \in \arg\max_{p_b|p_b^\dagger} \Big[(1-\alpha)\pi_b(p_b, \rho^*) + \alpha\pi_b\big(p_b, \sigma(p_b)\big)\Big]. \tag{A.45}$$

Using backward induction, firm a chooses $p_b^\dagger$ to yield its highest discounted payoffs that arise in this subgame.

Model (d): Algorithmic Stackelberg Leadership. At a reprogramming date, firm a chooses the algorithm parameters (ρ, σ) first; firm b observes the resulting algorithm and then chooses p_{bt}. In this sequential reprogramming-period timing, both price and the full form of the (A.40) is publicly observed before p_{bt} is chosen, and firm b solves

$$p_{bt}^* \in \arg\max_{p_b|(\rho, p_b^\dagger)} \Big[(1-\alpha)\pi_b(p_b, \rho) + \alpha\pi_b\big(p_b, \sigma(p_b)\big)\Big]. \tag{A.46}$$

If the target satisfies (A.42), then $p_{bt}^* = p_b^\dagger$ is optimal; otherwise firm b deviates immediately.

Table A-1 summarizes the equilibrium for the baseline timing and the four alternatives at various technology configurations $(\alpha, \beta\gamma) \in \{0,1\}^2$. We use the same benchmark labels as in the main text: "Bertrand" denotes the static Nash outcome with simultaneous pricing, "a as leader" and "a as follower" denote the corresponding sequential-move (Stackelberg-in-prices) outcomes, and "maximal coercion" refers to the outcome associated with the static program (21). Because pricing in the initial reprogramming period can differ from continuation play, the table reports continuation targets whenever $\beta\gamma = 1$; when $\beta\gamma = 0$ there is no continuation (every period is a reprogramming period), so the table reports the equilibrium for the reprogramming period instead.

The models with alternative timing assumptions yield two conclusions. First, maximal coercion is robust to the details of the reprogramming-period ordering: across all timing variants, the same maximal-coercion target remains an equilibrium prediction under the maximal technology configuration $(\alpha, \beta\gamma) = (1, 1)$. Second, the comparative statics continue to be organized by (i) an enforcement technology (speed, α) and (ii) a commitment device. Model (a) delivers the same predictions as the baseline. When the reprogramming period temporarily disables within-period punishment (Model (b)), results are also similar, however speed without persistence no longer generates supracompetitive outcomes. In the baseline model, commitment is delivered by persistence of the algorithm through $\beta\gamma$. In Models (c) and (d), the reprogramming-period move order itself creates an additional commitment channel, allowing firm a to attain equilibrium outcomes when $\beta\gamma = 0$ that require $\beta\gamma = 1$ in our baseline. These

Table A-1: Equilibria under Alternative Timing Assumptions

Timing Structure	$(\alpha, \beta\gamma)$			
	(0,0) No Speed/ No Commit	(0,1) Commit Only	(1,0) Speed Only	(1,1) Speed & Commit
Baseline	Bertrand	a as Leader	a as Follower	Maximal Coercion
Model (a): Only Punish Initially	Bertrand	a as Leader	a as Follower	Maximal Coercion
Model (b): No Initial Punishment	Bertrand	a as Leader	Bertrand	Maximal Coercion
Model (c): a Sets Target First	Bertrand	a as Leader	Maximal Coercion	Maximal Coercion
Model (d): a Moves First	a as Leader	a as Leader	Maximal Coercion	Maximal Coercion

Notes: The table shows equilibrium outcomes under the baseline timing (Section 2) and the four alternative reprogramming-period timing structures described in Appendix A.2. The columns correspond to limiting values of the algorithmic firm's technology parameters $(\alpha, \beta\gamma)$: $(0,0)$ represents no speed advantage or commitment across periods; $(0,1)$ represents full commitment without speed; $(1,0)$ represents maximum speed without commitment; and $(1,1)$ represents the combination of maximum speed and full commitment. In columns with $\beta\gamma = 1$, the entry reports the equilibrium target pair $(p_a^\dagger, p_b^\dagger)$ implemented in periods $t > \tau$ when the algorithm remains in force. In columns with $\beta\gamma = 0$ the algorithm is reprogrammed each period, so the entry instead reports the equilibrium outcome of the reprogramming-period pricing game implied by the corresponding timing structure.

timing variants therefore clarify that the mechanism is not a knife-edge consequence of a particular within-period ordering, but rather reflects the interaction of enforcement capacity and a commitment channel.

Existence and uniqueness of equilibria under these timing assumptions require the standard single-peakedness and best-response slope conditions that are met in standard demand environments and imposed as our regularity assumptions R1(i) and R1(ii). None of these timing variants require R1(iii) to ensure the objective function are well-behaved. Note that timing variants that endow firm a with an explicit first-mover commitment channel (Models (c) and (d)) can be characterized by standard backward induction: firm b observes the rule and then either complies with the target or deviates, and firm a chooses the rule anticipating that best response. In that environment, the regularity assumption R1(iii) (maximal incentive-compatible target), is naturally interpreted as the optimal choice from firm a's maximization problem subject to firm b's incentive constraint.

A.3 Extension to Oligopoly

While we primarily consider the case where an algorithmic firm faces a single slower rival, it is straightforward to extend the results to a more general setting in which a single algorithmic firm faces multiple slower rivals that all have the same pricing frequency. Thus, the speed advantage α characterizes the reaction time that the faster firm has relative to each of the slower firms.

We consider an n firm oligopoly with one algorithmic firm and $n-1$ slower rivals. The algorithmic firm's price is given by p_a while the vector of prices for slower rivals is given by $\mathbf{r} = (r_1, r_2, \ldots, r_{n-1})$. Generalizing the model in Section 4, firm a chooses the target price for itself, $p_a^\dagger$, and rivals, $\mathbf{r}^\dagger = (r_1^\dagger, r_2^\dagger, \ldots, r_{n-1}^\dagger)$, that maximize its discounted profits, subject to the algorithm technology and the incentive compatibility constraint for the rival firms:

$$\max_{(p_a^\dagger, \mathbf{r}^\dagger)|\mathbf{r}_t} (1-\alpha)\pi_a(\rho, \mathbf{r}_t) + \alpha\pi_a(\sigma(\mathbf{r}_t), \mathbf{r}_t) + \frac{\beta\gamma}{1-\beta\gamma}\pi_a(p_a^\dagger, \mathbf{r}^\dagger) \tag{A.47}$$

$$\text{s.t.} \quad (i) \quad \rho = p_a^\dagger \tag{A.48}$$

$$(ii) \quad \sigma(\mathbf{r}) = \begin{cases} p_a^\dagger & \text{if } \mathbf{r} = \mathbf{r}^\dagger \\ R_a(\mathbf{r}) & \text{if } \mathbf{r} \neq \mathbf{r}^\dagger \end{cases} \tag{A.49}$$

$$(iii) \quad \pi_j(p_a^\dagger, \mathbf{r}^\dagger) \geq (1-\alpha)\pi_j(\rho, \hat{r}_j, \mathbf{r}_{-j}^\dagger) + \alpha\pi_j(\sigma(\hat{r}_j, \mathbf{r}_{-j}^\dagger), \hat{r}_j, \mathbf{r}_{-j}^\dagger) \quad \forall \hat{r}_j,\ \forall j \tag{A.50}$$

After a fraction $1-\alpha$ of period t elapses, the algorithmic firm observes the vector of prices $\mathbf{r_t}$ and maintains price $p_{at}^\dagger$ if all prices are equal to the target prices for rivals. Otherwise, the algorithmic firm best responds to rivals, setting price $R_a(\mathbf{r})$.

We examine the case of linear demand that generalizes the duopoly demand given by equation (22). Demand for the algorithmic firm is given by

$$D_a(p_a, \mathbf{r}) = 1 - \left(\frac{1}{4} + \frac{d}{2}\right) p_a + \sum_{j=1}^{n-1} \frac{d}{2(n-1)} r_j \tag{A.51}$$

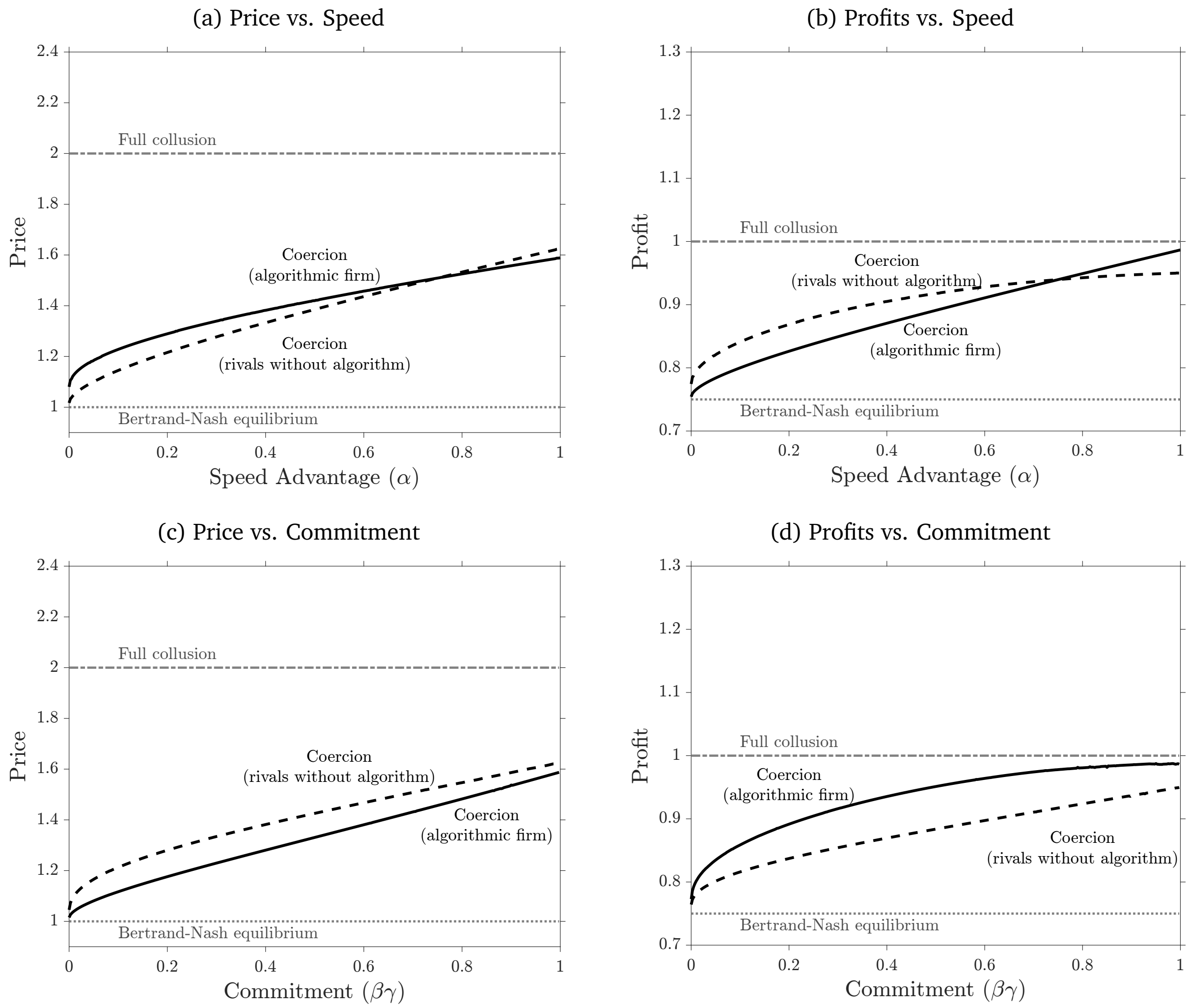
Demand is symmetric. Therefore, demand for rival j is given by

$$D_j(p_a, \mathbf{r}) = 1 - \left(\frac{1}{4} + \frac{d}{2}\right) r_j + \frac{d}{2(n-1)} p_a + \sum_{k \neq j} \frac{d}{2(n-1)} r_k \tag{A.52}$$

When $d = 1$, the Bertrand price and joint profit maximization prices are the same as for the duopoly linear demand given by equation (22).

We simulate equilibrium outcomes for the $n = 3$ case with $d = 1$. Appendix Figure A-1 shows equilibrium prices (panel (a)) and profits (panel (b)) with full commitment and different values for the speed advantage. The chart is qualitatively similar to the case of two firms shown in Figure 3. As in our previous analysis, prices are increasing in the speed advantage of firm a.

Figure A-1: Prices and Profits in Coercive Equilibrium with Three Firms

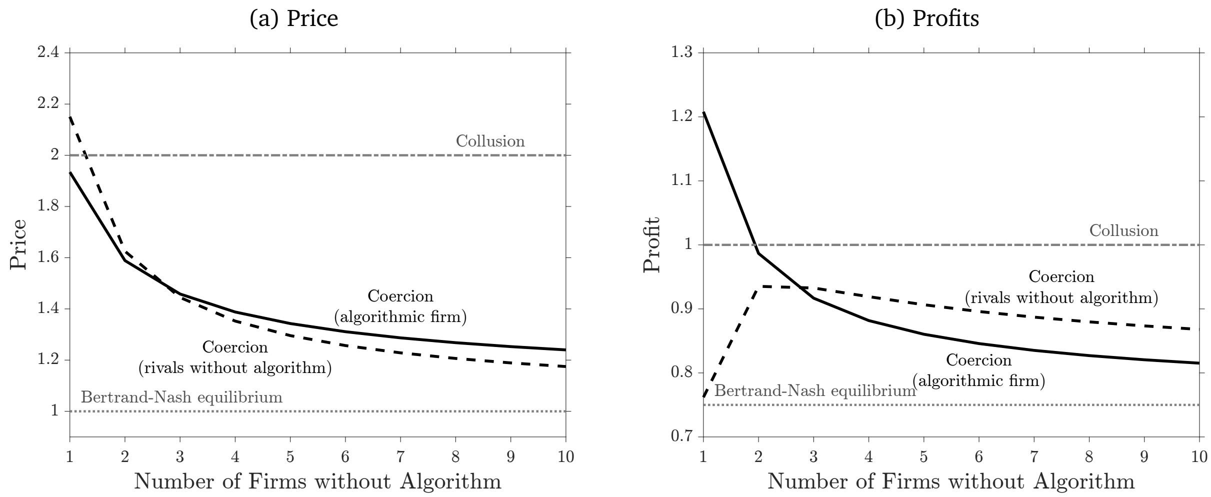


Notes: Panels (a) and (b) show equilibrium prices and profits for different values of speed advantage α under coercion with full commitment ($\beta\gamma = 1$). Panels (c) and (d) show equilibrium prices and profits under coercion for different values of commitment, $\beta\gamma$, when the faster firm has maximum speed ($\alpha = 1$). Outcomes under Bertrand competition with simultaneous pricing and joint profit maximization are displayed for comparison. Assumes $d = 1$ under the linear demand system given by equation (A.51).

Profits of firm a are also increasing in the speed advantage. However, the ability of firm a to coerce rivals into setting higher prices is somewhat muted compared to the two-firm case due to the fact that the incentive compatibility constraints are more binding in the three-firm case.

Panels (c) and (d) of Appendix Figure A-1 show equilibrium prices and profits in the case with maximum speed advantage and different values for the degree of commitment. Again, the results are similar to Figure 3. Prices and profits are increasing in the degree of commitment; however, prices under full commitment are lower than in Figure 3 given the additional constraints on firm a. In the two-firm case, the incentive compatibility constraint for the firm

Figure A-2: Prices and Profits in Maximal Coercive Equilibrium under Oligopoly



Notes: Panel (a) shows prices and panel (b) shows profits under Bertrand competition with simultaneous pricing, the coercive equilibrium with maximum speed ($\alpha = 1$) and maximum commitment ($\beta\gamma = 1$), and joint profit maximization. Assumes $d = 1$ under the linear demand system given by equation (A.51).

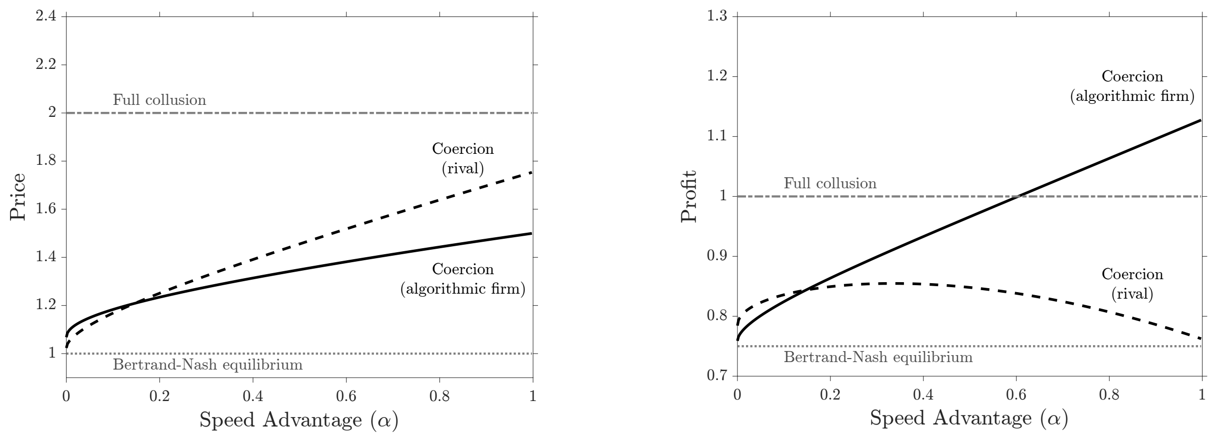
without the pricing algorithm is determined by $(1-\alpha)\pi_b(\hat{p}_b, \rho) + \alpha\pi_b(\hat{p}_b, \sigma(\hat{p}_b))$, which does not depend on the degree of commitment. However, in the three-firm case, the equivalent expression, $\pi_b(p_a^\dagger, r_b^\dagger, r_c^\dagger) \geq (1-\alpha)\pi_b(\rho, \hat{r}_b, r_c^\dagger) + \alpha\pi_b(\sigma(\hat{r}_b, r_c^\dagger), \hat{r}_b, r_c^\dagger)$, does depend on the degree of commitment since commitment affects the other slow firm's price, $r_c^\dagger$.

Finally, we examine the maximal coercion equilibrium by the number of firms in Appendix Figure A-2. Given the demand system, the Bertrand price and profits under joint profit maximization are constant as the number of firms increases. Under the coercive equilibrium, prices are decreasing as the algorithmic firm faces additional rivals without an algorithm. The algorithmic firm's profits are also decreasing in the number of rivals. However, the effect of additional rivals without an algorithm is relatively modest. Prices and profits are still substantially higher relative to the Bertrand equilibrium even with several rivals. With 10 rival firms, market prices are on average 18 percent higher than the Bertrand equilibrium.

A.4 Additional Tables and Figures

Figure A-3: Prices and Profits in Coercive Equilibrium: Partial Commitment

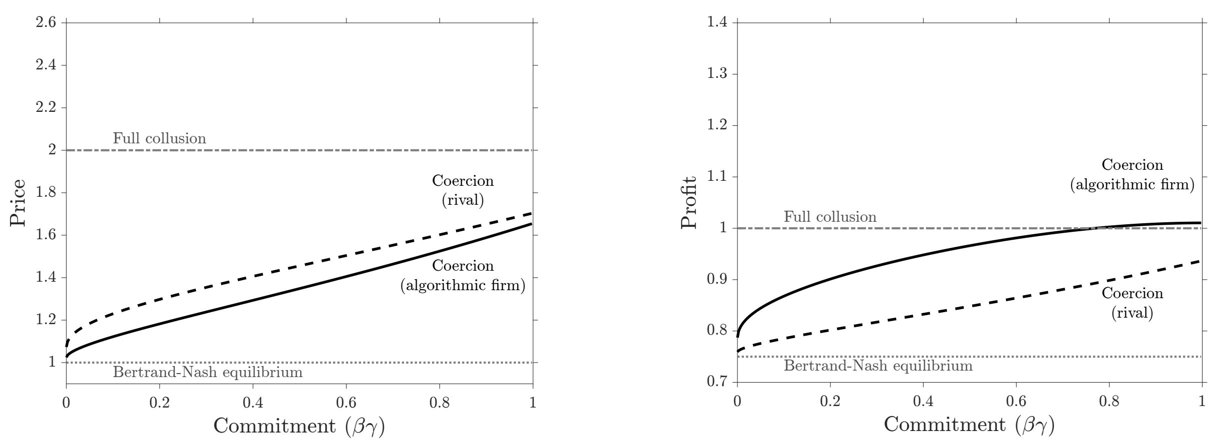
(a) Price (b) Profits



Notes: Panel (a) shows prices and panel (b) shows profits under Bertrand competition with simultaneous pricing, the coercive equilibrium with partial commitment ($\beta\gamma = 0.5$) and speed advantage α, and joint profit maximization. Assumes $d = 1$ under the linear demand system given by equation (22).

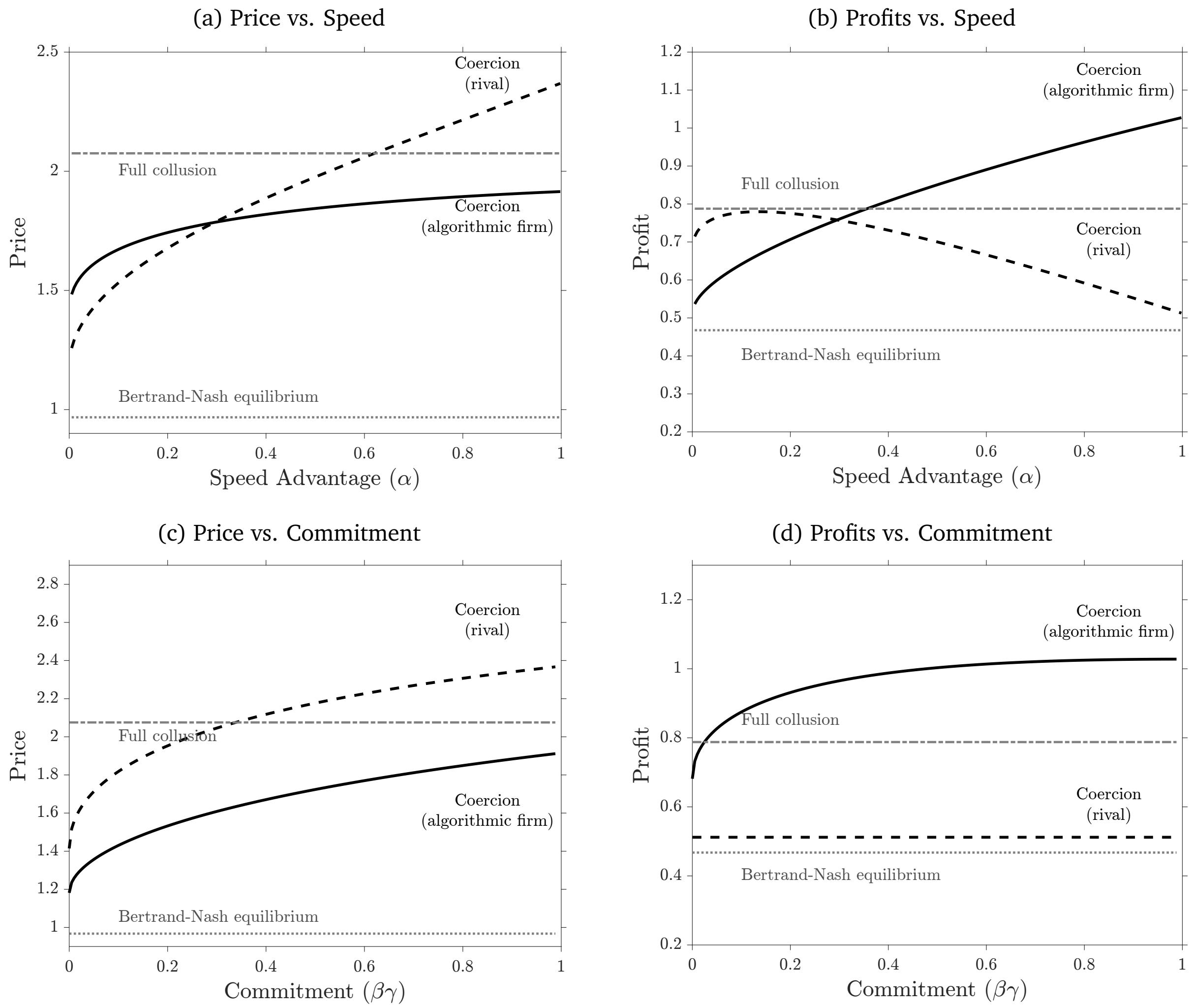
Figure A-4: Prices and Profits in Coercive Equilibrium: Intermediate Speed Advantage

(a) Price (b) Profits



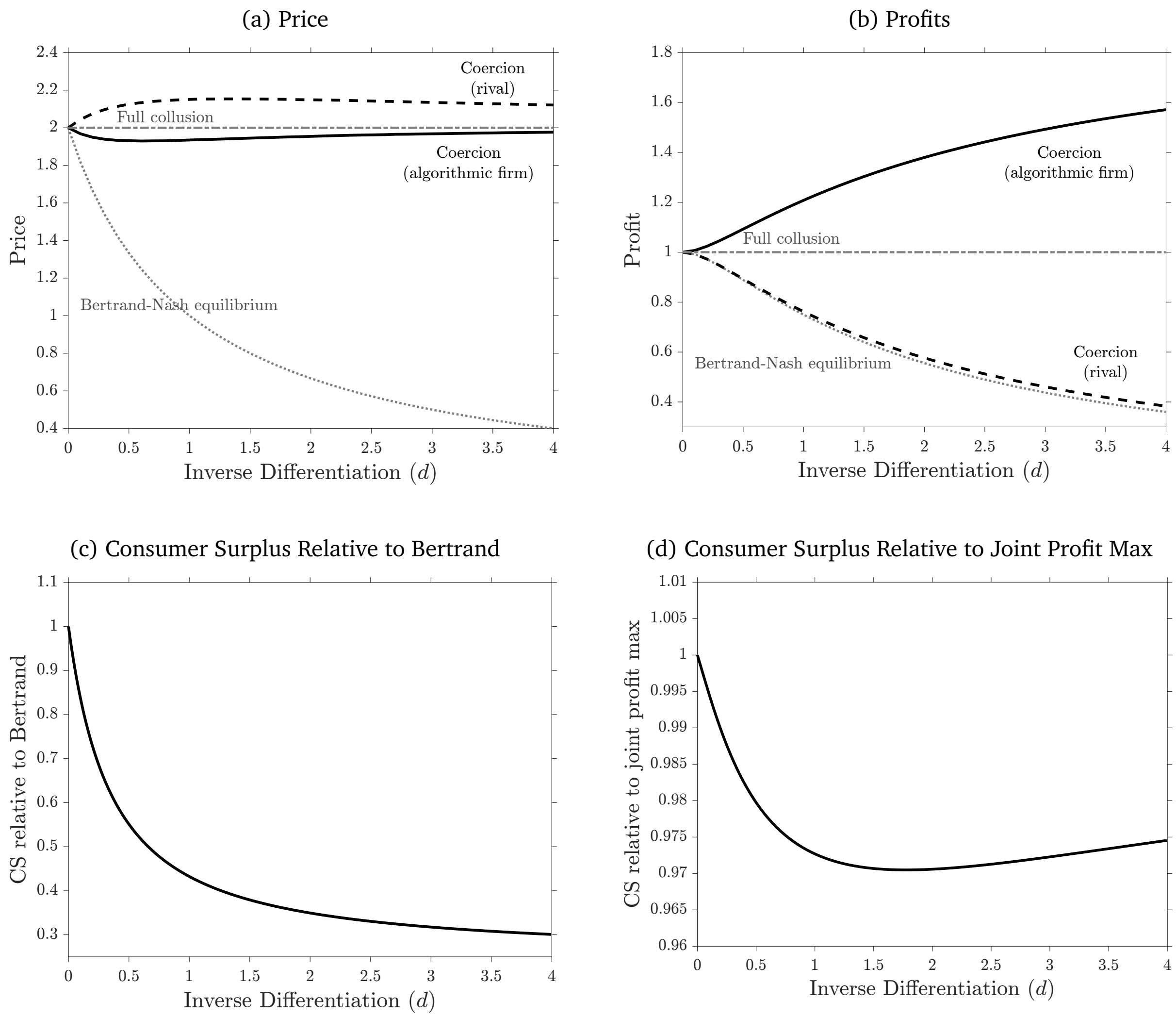
Notes: Panel (a) shows prices and panel (b) shows profits under Bertrand competition with simultaneous pricing, the coercive equilibrium with $\alpha = 0.5$ and commitment $\beta\gamma$, and joint profit maximization. Assumes $d = 1$ under the linear demand system given by equation (22).

Figure A-5: Prices and Profits in Coercive Equilibrium: Logit Demand



Notes: Panels (a) and (b) show equilibrium prices and profits for different values of the speed advantage α under the coercive equilibrium with full commitment ($\beta\gamma = 1$). Panels (c) and (d) show equilibrium prices and profits under coercion for different values of commitment, $\beta\gamma$, when the faster firm has maximum speed ($\alpha = 1$). Outcomes under Bertrand competition with simultaneous pricing and joint profit maximization are displayed for comparison. Assumes logit demand given by $D_i(p_i, p_{-i}) = \exp(-\eta p_i)/[\zeta + \exp(-\eta p_i) + \exp(-\eta p_{-i})]$ where $\eta = 2$ and $\zeta = 0.01$.

Figure A-6: Prices and Profits in Coercive Equilibrium, by Product Differentiation



Notes: Panel (a) shows prices and panel (b) shows profits under Bertrand competition with simultaneous pricing, coercion, and joint profit maximization. Panel (c) shows consumer surplus under the coercive equilibrium relative to Bertrand and panel (d) shows consumer surplus under the coercive equilibrium relative to joint profit maximization. Considers the maximal coercion case with $\alpha = 1$ and $\beta\gamma = 1$. Figures show equilibrium for different differentiation parameters d using linear demand given by equation (22).

Table A-2: Periods to Convergence for Coercive Equilibrium with Linear Pricing Rule

(a) Convergence Criterion of 1%

λ	Mean	Min	10th Percentile	25th Percentile	50th Percentile	75th Percentile	90th Percentile	Max
0.2	27.8	3	16	23	30	35	36	37
0.4	13.1	3	8	11	14	16	17	17
0.6	7.7	3	6	7	8	9	9	9
0.8	6.7	3	5	5	8	8	9	9
1	7.1	3	4	7	7	8	8	10
1.2	8.3	3	6	7	9	10	10	10

(b) Convergence Criterion of 0.1%

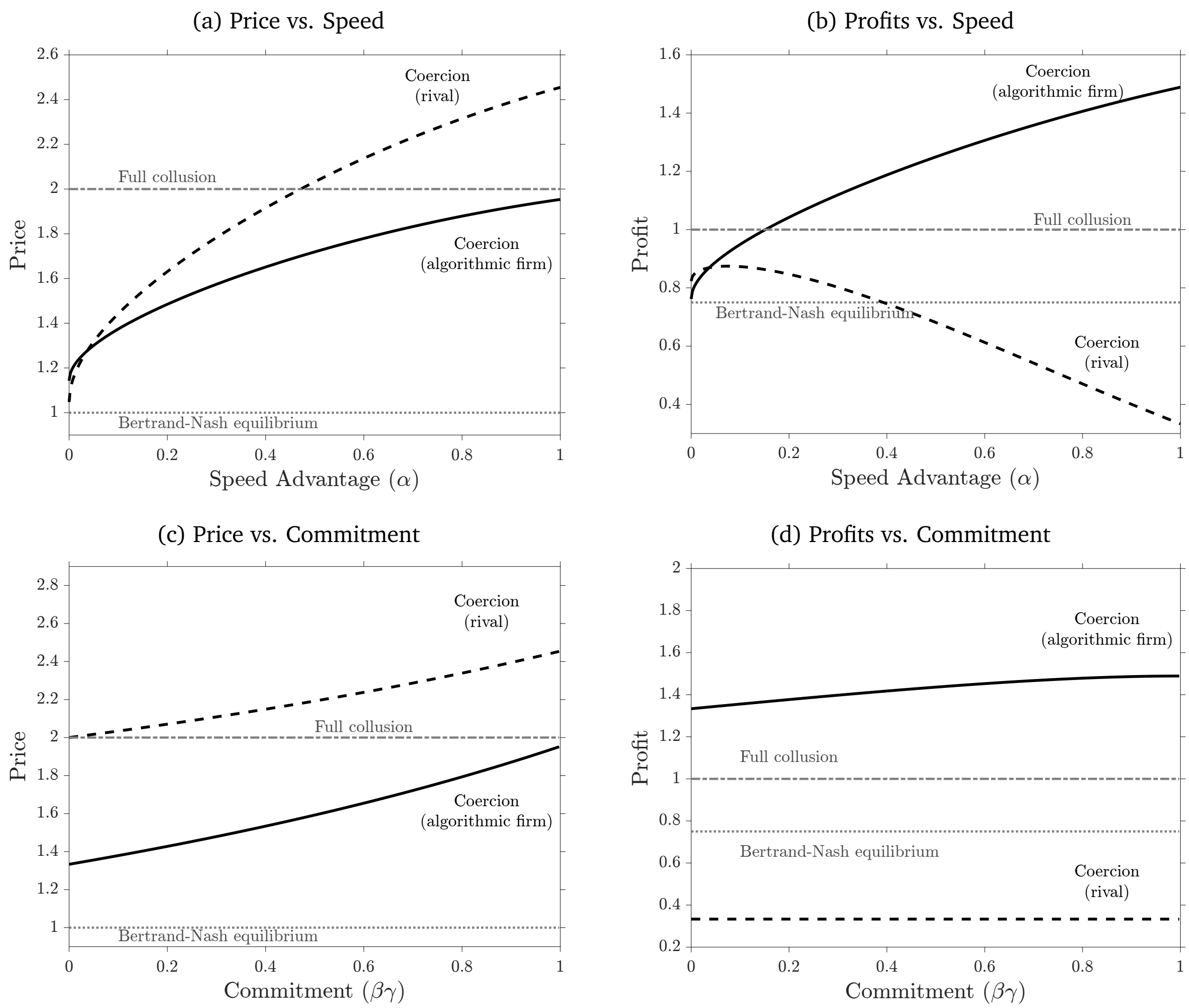
λ	Mean	Min	10th Percentile	25th Percentile	50th Percentile	75th Percentile	90th Percentile	Max
0.2	45.2	3	33	40	48	52	54	55
0.4	20.2	3	15	18	21	23	24	24
0.6	10.6	3	9	10	11	12	12	12
0.8	9.4	6	8	9	10	10	10	10
1	10.6	6	8	10	11	11	12	14
1.2	12.4	7	10	12	13	13	15	15

(c) Convergence Criterion of 0.01%

λ	Mean	Min	10th Percentile	25th Percentile	50th Percentile	75th Percentile	90th Percentile	Max
0.2	62.6	3	51	58	65	70	71	72
0.4	27.4	4	23	25	28	30	31	31
0.6	13.4	6	12	13	14	14	15	15
0.8	12.9	8	10	13	14	14	14	14
1	14.3	9	12	14	15	15	15	17
1.2	16.8	11	13	16	16	19	19	19

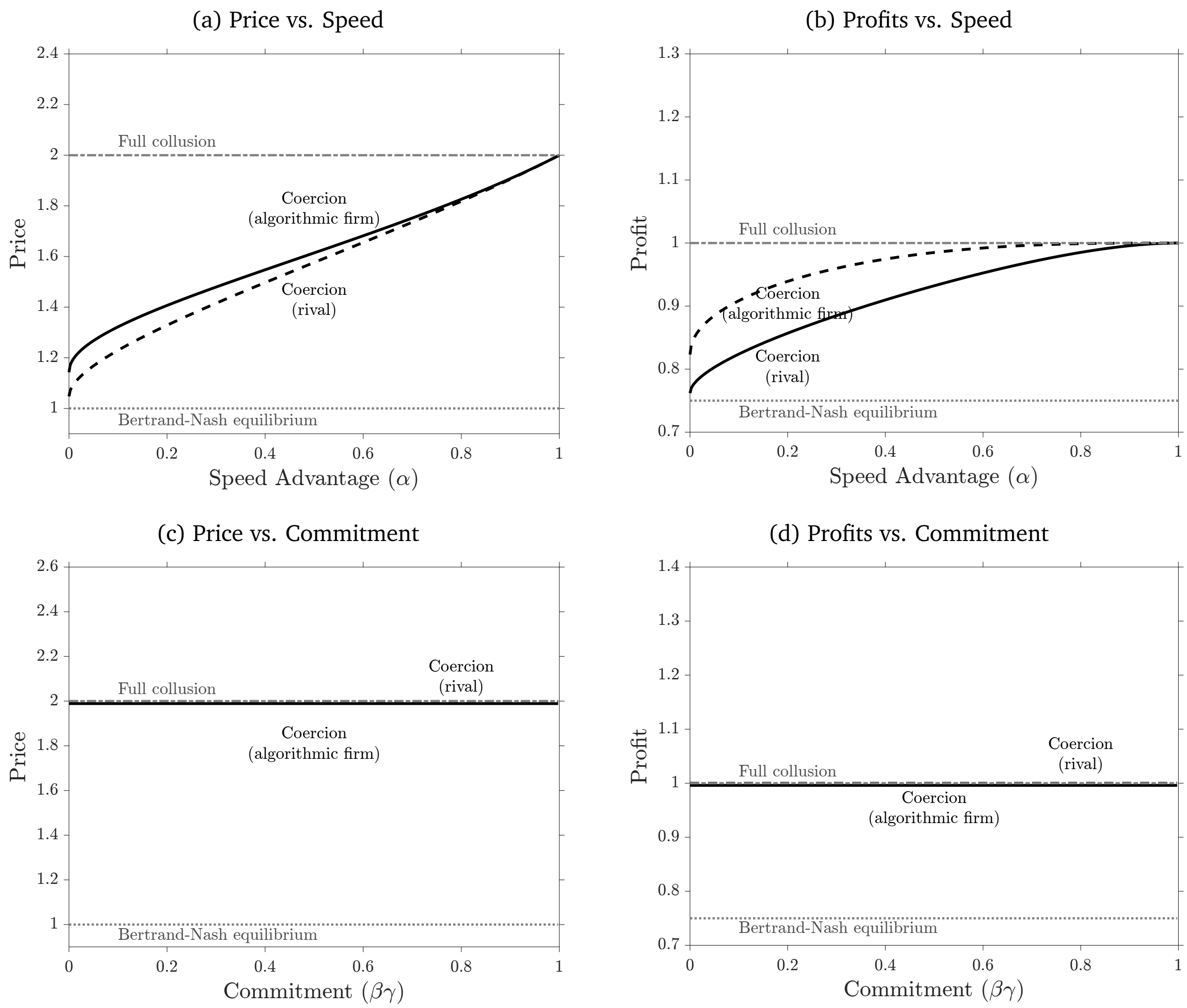
Notes: Table shows statistics of the number of periods until convergence when the algorithmic firm uses a linear pricing rule and the rival is assumed to use gradient learning. Rows indicate values of the price adjustment step size, λ. Convergence is defined as being within 1% of the equilibrium price for panel (a), within 0.1% of the equilibrium price for panel (b), and within 0.01% of the equilibrium price for panel (c). For each value of λ, we simulate price paths from 1,000 values of initial starting prices from a grid between 0 and 2 and report the mean, min, max, and percentiles. Simulation assumes algorithmic firm has maximum pricing speed ($\alpha = 1$) and uses the linear pricing rule described in Section 5.2. Assumes $d = 1$ under the linear demand system given by equation (22).

Figure A-7: Prices and Profits in Coercive Equilibrium: Harsh Punishment



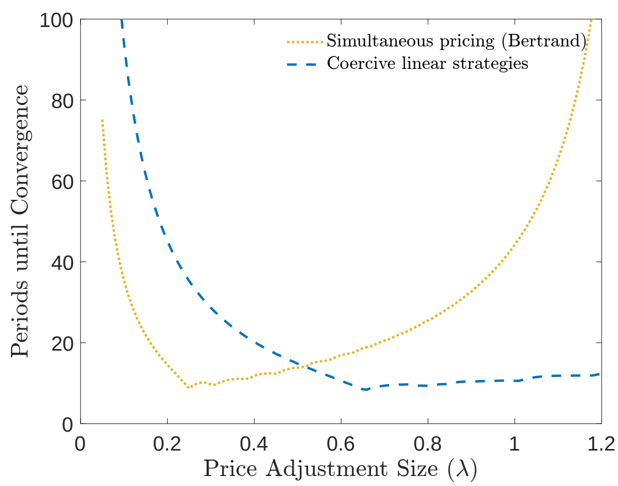
Notes: Panels (a) and (b) show equilibrium prices and profits for different values of pricing speed α under coercion with full commitment ($\beta\gamma = 1$). Panels (c) and (d) show equilibrium prices and profits under coercion for different values of commitment, $\beta\gamma$, when the faster firm has maximum speed ($\alpha = 1$). Outcomes under Bertrand competition with simultaneous pricing and joint profit maximization are displayed for comparison. Unlike Figures 3 and 4, we assume firm a uses a punishment price of 0 rather than the static best response. Assumes $d = 1$ under the linear demand system given by equation (22).

Figure A-8: Prices and Profits in Coercive Equilibrium: Price Matching Punishment



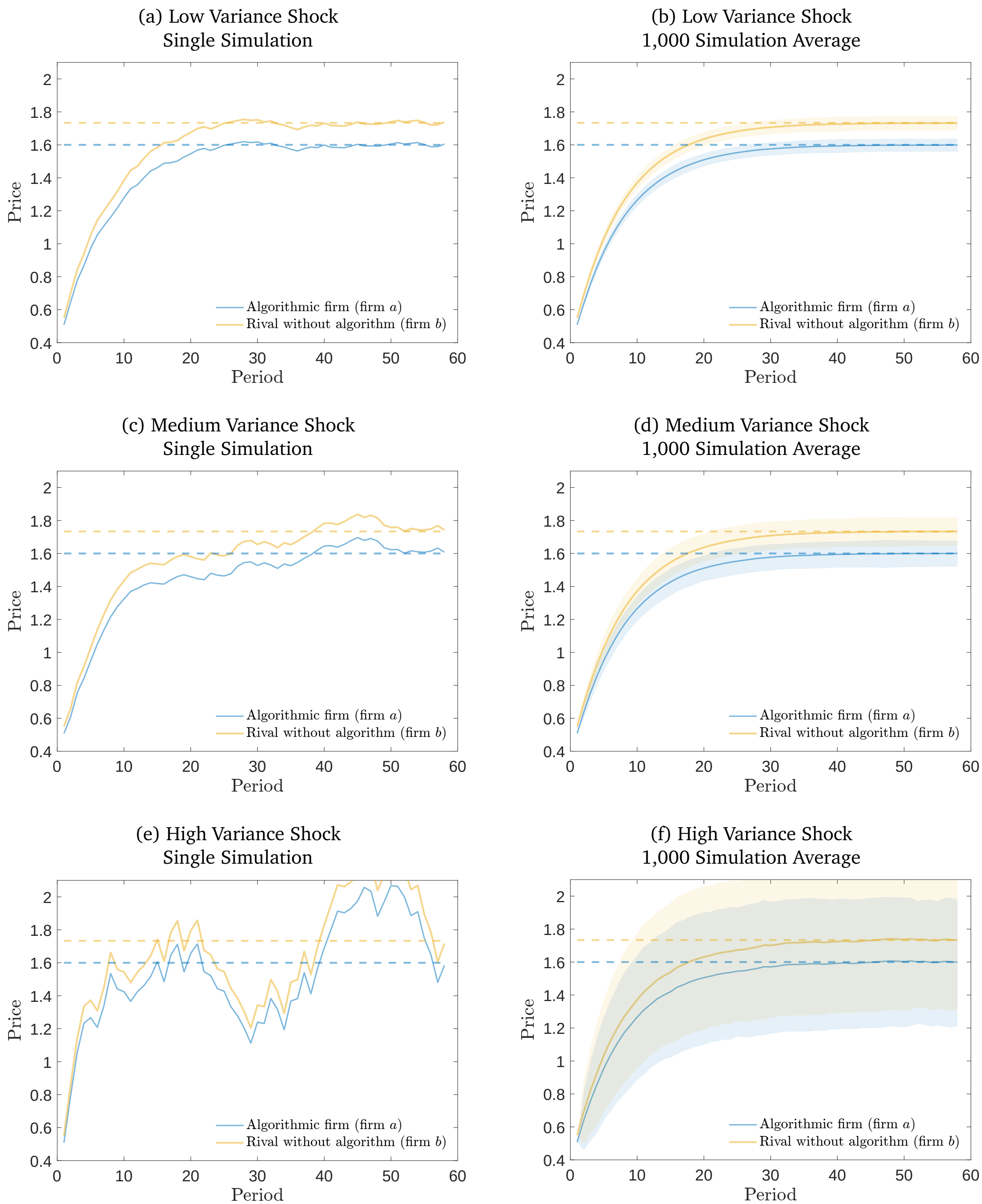
Notes: Panels (a) and (b) show equilibrium prices and profits for different values of speed advantage α under coercion with full commitment ($\beta\gamma = 1$). Panels (c) and (d) show equilibrium prices and profits under coercion for different values of commitment, $\beta\gamma$, when the faster firm has maximum speed ($\alpha = 1$). Outcomes under Bertrand competition with simultaneous pricing and joint profit maximization are displayed for comparison. Unlike Figures 3 and 4, we assume firm a uses a punishment function that matches rival's price rather than the static best response. Assumes $d = 1$ under the linear demand system given by equation (22).

Figure A-9: Convergence Speed with Gradient Learning by Price Step Size



Notes: The figure shows the mean number of periods until convergence when a naive firm uses gradient learning with price adjustment step size λ. Convergence is defined as being within 0.1% of the equilibrium price. For each value of λ, we simulate price paths from 1,000 values of initial starting prices from a grid between 0 and 2 and then average the number of periods until convergence. The orange line shows a simultaneous pricing game that results in Bertrand prices. The blue line shows the game in which the algorithmic firm has maximum pricing speed ($\alpha = 1$) and uses the linear pricing rule described in Section 5.2. Assumes $d = 1$ under the linear demand system given by equation (22).

Figure A-10: Convergence of Simulated Learning with Demand Shocks



Notes: The charts show simulated price paths when firm b uses gradient learning with uniformly distributed demand shocks with mean 1. Top panel shows shock with standard deviation 0.05, middle panel shows shock with standard deviation 0.1, and bottom panel shows shock with standard deviation 0.5. Demand shock affects the constant in linear demand given by equation (22). Left charts show a single simulation while right charts show the average of 1,000 simulations with error bars showing two standard deviations from the mean. Dashed lines show equilibrium prices derived in the proof for Proposition 8. The figure considers the case with $\alpha = 1$ and $\beta\gamma = 1$.