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 August 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, Federal Trade Commission seminar, UAB/IAE-CSIC seminar, CEPR-JIE/AI Conference, Algorithms and Markets Symposium, BECCLE Competition Policy Conference, Chicago Booth Market Design Conference, and Bates White seminar. 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.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2025 by Zach Y. Brown and Alexander MacKay. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Algorithmic Coercion with Faster Pricing
Zach Y. Brown and Alexander MacKay
NBER Working Paper No. 34070
July 2025, Revised August 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 a coercive outcome. 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. This framework, in which all firms coordinate on dynamic punishment strategies, provides the lens through which researchers and policymakers typically view supracompetitive prices.[1]

However, with pricing algorithms, other forms of supracompetitive pricing are possible. 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., Assad et al., 2024; Aparicio et al., 2024; Brown and MacKay, 2023). Automated responses to rival prices and faster pricing are defining features of algorithms that have drawn the attention of antitrust authorities.[2] More broadly, the growing use of algorithms and the possibility that they enable new pricing strategies have raised concerns among policymakers worldwide.[3]

In this paper, we investigate the extent to which a single firm with a pricing algorithm can induce supracompetitive prices, even when rivals maximize static profits and cannot collude. We call this form of conduct *coercion*. In our model, the pricing algorithm gives one firm the ability to adjust prices more frequently and (imperfect) commitment to a pricing rule across periods. The speed advantage lets the algorithmic firm reshape the rival's current-period incentives, while persistence encourages the algorithmic firm to not chase short-run gains. In equilibrium, the algorithmic firm can single-handedly implement prices substantially above competitive levels. In the limiting case, coercion can be worse for consumers than if firms explicitly agreed to set higher prices. Unlike collusion, coercion does not require coordination, mutual patience, or future punishments, nor does it require all firms to adopt similar algorithms.

The two technological features we study exist in a variety of markets. Online retailers use pricing technologies ranging from "real-time" to daily or weekly updates, and firms with 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] The U.K. Competition and Markets Authority has emphasized that pricing algorithms "almost instantly observe all competitors' prices, detect any deviation and implement a price response" (Competition and Markets Authority, 2018). In the U.S., the Federal Trade Commission alleges that Amazon's first-party algorithm can detect price changes for thousands of popular products within hours and respond automatically (Federal Trade Commission et al., 2023).

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

fastest technology react quickly to price changes by slower rivals (Brown and MacKay, 2023).[4] In retail gasoline, stations that adopt pricing algorithms have more frequent price changes than non-adopters (Assad et al., 2024). There are similar speed asymmetries in other markets, including short-term rental platforms, hotels, and, with the adoption of electronic shelf labels, brick-and-mortar grocery stores.[5] The pricing rules themselves can also be persistent relative to the prices they generate: an existing rule may update prices within minutes, while activating a new rule or changing its parameters can take longer.[6] A firm may adopt an algorithm and expect its pricing rule to be in place for months or even years.

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, due to technological or organizational factors. 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.

In a coercive equilibrium, 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 incentive compatibility constraint. In the baseline analysis, the algorithmic firm responds to deviations with its one-shot best response, while the rival myopically maximizes its current-period profit. The latter assumption rules out equilibria that rely on the threat of future punishment, as in standard models of collusion. Instead, coercion arises because the algorithm's speed advantage allows it to respond within the period, reshaping the perceived profits of its rival. High-speed pricing algorithms thus offer an alternative channel through which firms can raise prices.[7]

The model nests familiar benchmarks. If the algorithm offers no speed advantage but provides full commitment across periods, the algorithmic firm behaves as a leader in a sequential-

[4]Providers of pricing tools for online sellers explicitly advertise the importance of being faster than rivals. For instance, one provider states that businesses can "compete more effectively by responding quickly" to rivals' price changes (DealHub, 2025). Another prominent repricing service offers a basic version that updates prices hourly and a premium version that "reacts to changes your competitors make in 90 seconds" in order to "beat competitors with super-fast repricing" (Repricer.com, 2023).

[5]On short-term rental platforms, many hosts set prices manually, while professional hosts can use tools that automatically update nightly rates in response to rivals (Airbnb, 2026). Hotel chains and large properties use revenue-management systems that track competitors and reprice rooms intraday, whereas many independent hotels adjust prices manually. For example, Lighthouse advertises tools using real-time pricing data (Lighthouse, 2026). Electronic shelf labels likewise let grocers reprice far faster than conventional price tags (Stamatopoulos et al., 2021).

[6]Online retailers, such as Amazon and German platform Kaufland, have documentation stating that, while pricing rules can autonomously adjust prices quickly, there is a delay when activating or switching the pricing rule itself (Amazon, 2026; Kaufland Global Marketplace, 2026a,b).

[7]It is well known that, in simultaneous-move games, higher-frequency pricing implies a larger per-period discount factor, making collusion easier to sustain when all firms are forward-looking (e.g., Abreu et al., 1991).

move (Stackelberg pricing) game. Prices increase because the firm can commit to not chase short-run gains. If the algorithm can react instantaneously but does not enable commitment across periods, the algorithmic firm behaves instead as a follower that best responds to the rival.[8] Our focus is on the interaction of these two features. With both commitment and a speed advantage, the algorithmic firm can unilaterally implement substantially higher prices than under either benchmark.

Our results characterize equilibrium prices, profits, and welfare in coercive equilibrium. Even when the algorithmic firm's speed advantage and commitment are modest, it can induce prices that provide higher profits than it would receive as a sequential leader or follower. *Maximal coercion* is the limiting case in which the algorithm provides an instantaneous reaction and full commitment to a rule across periods. In this case, the algorithmic firm can obtain more than its full collusion profits by inducing the rival to charge a price above the fully collusive price while itself setting a slightly lower price. Average prices can exceed collusive prices, and consumer surplus and total welfare can fall below the full collusion benchmark. Thus, coercion can have substantial effects on equilibrium prices and welfare.

The analysis highlights the complementarity between speed and commitment. When both features are present, the algorithmic firm can discipline its rival through rapid punishment and discipline itself through commitment,[9] thereby allowing it to obtain higher profits than under either feature alone. The limiting case of maximal coercion has a mechanism-design flavor: the firm chooses its preferred target prices subject to its rival's incentive constraint. In this way, algorithms can increase prices without transfers, contracts, or explicit agreements.

We characterize when the algorithmic firm can implement the collusive outcome. Even when this outcome is feasible, it need not be selected in equilibrium. The algorithmic firm can often increase its own profit by choosing a different target price vector that continues to satisfy the rival's incentive compatibility constraint. Thus, coercion need not reproduce the collusive outcome even when collusive prices are implementable.

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. We therefore study an extension in which the rival 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. When the algorithmic firm commits to a linear pricing rule in the rival's price, the same learning process instead converges to a coercive outcome 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,

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

[9] Algorithm providers also emphasize this benefit of commitment, suggesting that persistent rules guard against "rash" or overly short-sighted pricing decisions (FeedbackWhiz, 2025; Informed Software, 2025).

thereby guiding the rival toward higher prices.[10] Convergence to coercive outcomes is rapid on average across several simple learning rules.

The framework also has implications for platform design. Large online platforms can influence how frequently sellers observe rivals' prices, how quickly sellers may update prices, and which repricing tools are available to them. Moreover, platforms can provide different pricing technologies to different sellers.[11] We consider these design decisions through the platform's optimal choice of speed and commitment. A platform that places sufficiently high weight on seller profits has an incentive to enable faster or more persistent algorithms for some sellers, even when doing so harms consumers. A platform that places sufficient weight on consumer surplus instead restricts these advantages. Thus, seemingly technical choices about pricing tools can either facilitate competition or enable supracompetitive prices through coercion.

The coercive mechanism is robust to alternative assumptions. If the rival is forward-looking, the coercive outcome remains an equilibrium, although additional equilibria may also exist. The maximal-coercion outcome can be achieved under alternative timing assumptions at reprogramming dates. Other punishment functions, including harsher punishments and price matching, also generate supracompetitive prices. With sufficiently fast pricing, price matching can sustain fully collusive prices, though the algorithmic firm may prefer best-response punishment because it shifts surplus in its favor. The results extend to an oligopoly setting, with one algorithmic firm and multiple slower rivals. Additional rivals make coercion harder because a deviating rival can steal from other slow firms whose prices do not respond, but a single fast and persistent pricing rule can still sustain supracompetitive outcomes.

Our results are developed for the case when one firm has superior access to monitoring or repricing technologies, consistent with the examples documented above. Our analysis also suggests why asymmetric pricing technologies may persist even when all firms can invest in faster repricing. A rival may not want to eliminate the speed advantage of an algorithmic firm. The asymmetry can soften price competition and allow both sellers to share in the additional profit. Hence, equal access to repricing tools need not imply equal use. This observation provides a rationale for studying asymmetric pricing technology and a motivation to focus on the potential for coercion.

Our analysis contributes to several strands of the literature. A growing theoretical literature studies how commitment by pricing algorithms affects 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. The role of algorithm speed has received less attention, even though rapid responses to rivals' prices are a defining feature of many pricing tools. One exception is Brown and MacKay (2023), who an-

[10] This mechanism resembles the strategic manipulation of beliefs in signal-jamming models of dynamic games (Fudenberg and Tirole, 1986).

[11] For example, in its Amazon Marketplace investigation, the European Commission found that Amazon Retail had more favorable access to pricing-related data than third-party sellers (European Commission, 2022a).

alyze Markov-perfect outcomes when firms differ in how frequently they can adjust prices.[12] In contrast, this paper is the first to explicitly analyze the implications of combining a speed advantage with multi-period commitment.

A parallel literature asks 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.[13] 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 learn to collude, we show how one firm can use speed and commitment to coerce a rival through current-period incentives.

As discussed above, the empirical literature documents that pricing technologies and response speeds vary across firms in a variety of markets (e.g., Brown and MacKay, 2023; Assad et al., 2024; Stamatopoulos et al., 2021). Byrne et al. (2025) link speed asymmetry to outcomes by studying a legal settlement that forced one gasoline retailer to exit a price-monitoring platform shared with its rivals. After the settlement, that firm updated prices less frequently and market prices rose. Other papers document the use of algorithmic pricing to enable automated rules on online platforms (Chen et al., 2016; Wieting and Sapi, 2021; Musolff, 2024).

Finally, the analysis relates to 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 and Wolitzky (2025) address issues that arise when multiple firms adopt third-party pricing tools or engage in information sharing. We instead focus on how platform design choices that influence speed and commitment shape competitive outcomes and platform incentives.

The paper proceeds as follows. We introduce the model and equilibrium in Section 2. In Section 3, we discuss benchmark cases. We examine prices, profits, and welfare in Section 4. We also explore private incentives for technology adoption, alternative punishments, and oligopoly. In Section 5, we introduce learning, and we show how the algorithmic firm can obtain coercive outcomes with simple linear rules. Section 6 addresses 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 its price directly. Time is continuous, but decisions are organized into discrete periods. The key

[12] Outside product markets, the high-frequency trading literature has examined incentives to invest in a speed advantage (Biais et al., 2015; Budish et al., 2015; Aquilina et al., 2022).

[13] 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.

asymmetry is that the algorithm provides one firm with a pricing rule that reacts quickly to the rival's price and persists across periods. 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 $s \in [0, \infty)$, and it is partitioned into discrete periods $t \in \{0, 1, 2, \ldots\}$. Period t corresponds to the 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 shifter of demand or costs observed at the start of each period. The equilibrium conditions and comparative statics below are conditional on the current state persisting.

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, $P_i = [0, p_i^{\max}]$. Each $p_i^{\max}$ is sufficiently large that it does not bind for any optimal response considered in the analysis. Since π_i is strictly concave in its own price, firm i's static best-response function $R_i(\cdot)$ exists, is single-valued, and is continuous.

The period length is such 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 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 $\beta \in [0, 1)$.[14] 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.

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,

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

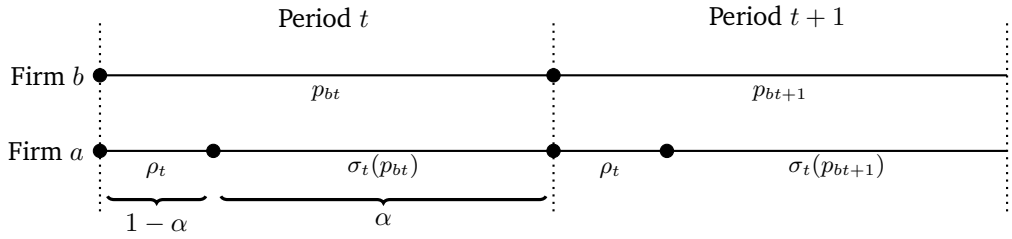
for a candidate p_{bt}, where s is the index of continuous time. We express the algorithm as a function of p_{bt} and not p_{bs} because firm b sets its price only once per period.

For example, suppose that firm b updates its price once per hour. If firm a updates its price every 15 minutes, then $\alpha = 3/4$, as firm a's update rule applies to three-quarters of the one-hour period. If firm a also sets prices only once per hour, then $\alpha = 0$, 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 (e.g., within seconds) after firm b chooses its price.

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. This can happen when software systems limit the frequency of pricing updates. 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' prices at different frequencies because of different investments in price-monitoring technology. 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.

[14] 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 a 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 from α to $\tilde{\alpha}$ is one-to-one.

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 $\mathcal{A}_t^{\alpha,\gamma}$ at the start of period t and remains committed to that algorithm in period $t+1$, which occurs with probability γ. The example depicts $\alpha = 3/4$. Black dots are opportunities to adjust prices.

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.[15] In practice, this parameter reflects the number of pricing periods that a particular algorithm is expected to remain in force. Consider the hourly-pricing example above. If firm a updates its algorithm once per week on average, then $\gamma = 167/168 \approx 0.99$.

Appendix A.8 provides additional example values for α and γ implied by quarterly, monthly, weekly, daily, and hourly pricing intervals.

Figure 1 illustrates how the algorithmic technology changes within-period timing relative to 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. In the figure, the chosen algorithm is assumed to persist into period $t + 1$, which occurs with probability γ.

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

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

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

[15] We assume an exogenous commitment parameter to capture technological or institutional constraints on reprogramming frequency. We then endogenize this choice in Section 6.

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

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

We denote the product of firm a's discount factor and its commitment parameter, $\beta\gamma$, as the algorithmic firm's *effective commitment*.

At a reprogramming date, the baseline model has ρ, σ, and p_{bt} chosen simultaneously. Appendix A.2 considers four alternative timing assumptions, including a one-period implementation delay for the update rule and firm a first declaring $p^\dagger_b$ before ρ and p_{bt} are chosen. The coercive mechanism and the maximal-coercion outcome are robust across these alternative move orders.

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 best responds to $\mathcal{A}_\tau$ in the current period but ignores any implications for future periods. This assumption 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.[16]

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

[16] 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).

(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 conservative benchmark. In standard repeated games, the sustainability of supracompetitive prices relies on rivals valuing future profits enough to fear punishment. A myopic rival places no weight on continuation incentives. Consequently, the algorithmic firm cannot rely on the rival's patience to sustain high prices. Instead, it must rely on its speed advantage to make deviations unprofitable. Assumption A1 is not essential to the mechanism. Rather, it highlights that coercion can arise even in settings where collusion is not possible.

The baseline equilibrium remains an equilibrium when the rival is forward-looking and arbitrary history-dependent strategies are available. To see this, note that firm b's price does not affect the evolution of demand or costs, the persistence of the algorithm, or, given the baseline equilibrium algorithm of firm a, the decisions of firm a and the prices chosen by its algorithm in future periods. Firm b's continuation value is therefore independent of its current price, so it solves the same current-period problem as equation (8). Thus, no history-dependent deviation is profitable for firm b. Firm a then faces the same optimization problem as in the baseline model, and the equilibrium remains. Allowing history-dependent strategies can generate additional equilibria, but it does not eliminate the baseline.

Our baseline model assumes that the rival has complete information, understands the update rule, and chooses a best response. Section 5 relaxes the assumption by considering a naive rival that does not know the algorithm and observes only its own prices and realized profits.

2.4 Coercive Equilibrium

We now define the equilibrium concept. An equilibrium consists of an algorithm for firm a and a pricing strategy for firm b such that each is optimal given the other.

Definition 1 (Coercive Equilibrium)**.** *A coercive equilibrium is a strategy profile*

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

where $\mathcal{A}_\tau^* = (\rho_\tau^*, \sigma_\tau^*)$ denotes the algorithm chosen at reprogramming date τ and $\{p_{bt}^*\}$ is the sequence of prices chosen by firm b, such that:

(i) Given p_{bt}^*, each $\mathcal{A}_\tau^*$ solves firm a's dynamic program in (4)–(5) *at $t=\tau$, anticipating that firm b best responds to $\mathcal{A}_\tau^*$ in all continuation periods in which it remains in force; and*

(ii) Given the algorithm $\mathcal{A}_\tau^*$ in force at t, each p_{bt}^* solves firm b's static problem in (8)*.*

These conditions require that $(\mathcal{A}_\tau^*, p_{b\tau}^*)$ form a Nash equilibrium at each reprogramming date τ. With x_t constant, strategies are stationary, so $\mathcal{A}_\tau^* = \mathcal{A}^*$ for all τ and $p_{bt}^* = p_b^*$ for all t. We therefore describe equilibrium in terms of a single algorithm $\mathcal{A}^*$ and a single price p_b^*.

We focus on an economically transparent subclass of *target-price equilibria*. In this class, the algorithm maintains a constant within-period price, $p_a^\dagger$, if firm b chooses a price equal to a target price, $p_b^\dagger$, chosen by firm a. If firm b does not comply with the target price, the algorithm updates firm a's price after $(1-\alpha)$ of the period has elapsed. For the baseline analysis, we assume the algorithm punishes deviations from the target price using firm a's static best response.

Assumption A2 (Best-Response Punishment)**.** *For a target price $p_b^\dagger$, $\sigma(p_b^\dagger) = \rho$. Otherwise, the algorithm uses the static best response, i.e., $\sigma(p_b) = R_a(p_b)$ for any $p_b \neq p_b^\dagger$.*

Using the static best response is an economically natural update rule, as it delivers the highest payoff conditional on the price of firm b. The assumption, along with regularity conditions we discuss below, selects a particular class of equilibria. Assumption A2 is conservative relative to harsher punishment rules, such as marginal-cost pricing, which we address in Section 4.3.

Under Assumptions A1 and A2, it is convenient to re-express the algorithmic firm's problem from equation (7) in terms of target prices. Let $(p_a^\dagger, p_b^\dagger)$ denote the prices that firm a would like to be chosen in every period. The algorithm implements these prices along the equilibrium path and applies the selected punishment after off-path deviations. Given the current rival price p_{bt}, firm a chooses a target price pair and algorithm components (ρ, σ) that solve

$$\tilde{V}(t) = \max_{(p_a^\dagger, p_b^\dagger)} \Big[(1-\alpha)\,\pi_a(\rho, p_{bt}) + \alpha\,\pi_a(\sigma(p_{bt}), p_{bt}) + \frac{\beta\gamma}{1-\beta\gamma}\,\pi_a(p_a^\dagger, p_b^\dagger) \Big] \tag{10}$$

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

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

$$(iii) \quad \pi_b(p_b^\dagger, p_a^\dagger) \;\geq\; (1-\alpha)\,\pi_b(\hat{p}_b, \rho) + \alpha\,\pi_b\big(\hat{p}_b, \sigma(\hat{p}_b)\big) \quad \forall \hat{p}_b \neq p_b^\dagger. \tag{13}$$

Constraint (i) requires the initialization rule to implement firm a's target price. Constraint (ii) states that the update rule follows the same price when firm b complies and uses best-response punishment when it deviates. Constraint (iii) is the incentive compatibility condition: target prices must deliver at least as much profit to firm b as any deviation, given the algorithm's response.

We adopt a standard tie-breaking convention: when firm b is indifferent, it complies with the target. We also assume that, when firm a chooses among incentive-compatible targets, it does not consider its choice of $p_b^\dagger$ to affect firm b's compliance in the current period. We formalize this and other regularity conditions as Appendix Assumption R1, which we discuss below.

Under these assumptions, we obtain existence and uniqueness of equilibria within the selected class.

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 conduct assumption that limits the scope for supracompetitive pricing. Assumption A2 anchors off-path play and selects the baseline equilibrium from the broader class of Markov pricing rules with alternative punishment functions. Appendix Assumption R1 collects three regularity conditions used by Proposition 1. The first two conditions ensure that firm a's objective is well-behaved and that the mapping $p_{bt} \mapsto p_b^\dagger(p_{bt})$ has a unique fixed point. The third condition states that, when firm a compares incentive-compatible targets, it holds fixed whether the current price p_{bt} is treated as compliant. This restriction rules out behavior in which firm a adjusts the target $p_b^\dagger$ downward slightly to force noncompliance, even when firm b desires to comply. We view this as a natural restriction because firm a does not observe firm b's contemporaneous price at the time it chooses the algorithm, and it ensures intuitive behavior for interior parameter values in our baseline model.

The coercive mechanism does not require these particular assumptions. As discussed in Section 2.3, the equilibrium remains if Assumption A1 is relaxed. We consider alternative punishments to Assumption A2 in Section 4.3. The third condition of Appendix Assumption R1 is not required under alternative timing conventions, as we discuss in Appendix A.2. It is also unnecessary under, e.g., linear or marginal-cost punishments.

For the remainder of the paper, unless otherwise stated, Assumption A1, Assumption A2, and Appendix Assumption R1 are maintained, and references to coercive equilibrium prices, profits, and comparative statics pertain to the class of coercive equilibria characterized in Proposition 1.

3 Coercive Equilibrium: Benchmark Cases

We begin with benchmark environments that isolate the two features of the algorithmic technology. We consider outcomes with no speed advantage and no commitment, only commitment, only a speed advantage, and the interaction of the two. We then highlight maximal coercion, in which the firm has full commitment and a maximal speed advantage. We discuss the mechanisms driving the equilibrium prices in each case, compare the outcomes to the joint-profit-maximizing (collusive) outcome, and provide a numerical example.

3.1 Effects of Speed and Commitment

We first focus on the boundary cases of the technology to isolate forces in the model. We characterize outcomes with respect to firm a's effective commitment, $\beta\gamma$, and we write $\beta\gamma = 1$ for the full-commitment limit, formally $\beta\gamma \to 1$.

The model nests familiar benchmarks. With no speed advantage and no commitment, firm a's algorithm is a static pricing rule. Given Assumption A1, this leads to the simultaneous pricing outcome. When the algorithm enables full commitment but no speed advantage, firm a prices as a Stackelberg leader. Conversely, with a maximal speed advantage but no commitment, it

prices as a Stackelberg follower, as in Brown and MacKay (2023). Each of these outcomes lies on one or both firms' static best-response functions.

When the algorithm combines speed and commitment, equilibrium prices need not satisfy either best-response restriction. In the *maximal coercion* case, with full commitment and a maximal speed advantage, prices are constrained only by firm b's incentive compatibility constraint.

The following lemma formalizes these benchmark outcomes.

Lemma 1 (Benchmark Cases)**.** *The benchmark coercive equilibria are as follows.*

1. ***Simultaneous pricing.*** *If $(\alpha, \beta\gamma) = (0, 0)$, the firms' problems reduce to $\max_{p_a} \pi_a(p_a, p_b)$ and $\max_{p_b} \pi_b(p_b, p_a)$, yielding the Bertrand–Nash outcome (p_a^B, p_b^B) defined by $p_a^B = R_a(p_b^B)$ and $p_b^B = R_b(p_a^B)$.*

2. ***Commitment only.*** *If $(\alpha, \beta\gamma) = (0, 1)$, firm a's problem reduces to $\max_{p_a} \pi_a\big(p_a, R_b(p_a)\big)$, and the equilibrium converges to the sequential outcome (p_a^L, p_b^F) in which firm a leads and firm b follows, where p_a^L solves firm a's problem and $p_b^F = R_b(p_a^L)$.*

3. ***Speed only.*** *If $(\alpha, \beta\gamma) = (1, 0)$, firm b's problem reduces to $\max_{p_b} \pi_b\big(p_b, R_a(p_b)\big)$ and firm a chooses $p_a^\dagger = R_a(p_b^\dagger)$, yielding the sequential outcome (p_a^F, p_b^L) in which firm b leads and firm a follows, with $p_a^F = R_a(p_b^L)$.*

4. ***Maximal coercion.*** *If $(\alpha, \beta\gamma) = (1, 1)$, the target-price problem reduces to*

$$\begin{aligned} \max_{(p_a^\dagger, p_b^\dagger)} \quad & \pi_a(p_a^\dagger, p_b^\dagger) \\ \textit{s.t.} \quad & \pi_b(p_b^\dagger, p_a^\dagger) \geq \max_{\hat{p}_b} \pi_b\big(\hat{p}_b, R_a(\hat{p}_b)\big). \end{aligned} \tag{14}$$

   Unlike the first three cases, neither equilibrium target price need lie on a firm's static best-response function.

The first part of Lemma 1 shows that the model nests the familiar Bertrand–Nash benchmark with simultaneous pricing. Firm a's algorithm cannot respond within the period and must be reset each period. Therefore, firm a simply chooses ρ to be the usual static best response to firm b's price. Firm b chooses the usual static best response to ρ. Thus, supracompetitive outcomes are not built into the economic environment of the model; they require either a speed advantage, commitment, or both.

The second and third parts of Lemma 1 isolate these two instruments and highlight the different incentives that appear in the value function in (10), which correspond to the simultaneous (Bertrand) incentive, the Stackelberg follower incentive, and the Stackelberg leader incentive. We discuss these cases in more detail.

First consider an algorithm that only enables commitment. Setting $\alpha = 0$ yields the problem

$$\max_{\rho} \Big[\pi_a(\rho, p_{bt}) + \frac{\beta\gamma}{1-\beta\gamma}\,\pi_a(\rho, R_b(\rho))\Big]. \tag{15}$$

The first term captures the short-run simultaneous-pricing incentive. The second captures the Stackelberg leader incentive from committing to ρ and inducing firm b to play $R_b(\rho)$ in future periods. As $\beta\gamma \to 1$, the second term dominates and the program reduces to a Stackelberg pricing game as described in Lemma 1 above. Under standard regularity conditions, if $\alpha = 0$ and $\beta\gamma \in (0, 1)$, the equilibrium price vector lies on firm b's static best-response function between the Bertrand outcome and (p_a^L, p_b^F). When prices are strategic complements, this commitment motive raises both prices above the Bertrand benchmark.

With a speed advantage but no commitment ($\beta\gamma = 0$), the continuation term disappears. The algorithmic firm's objective is instead

$$\max_{\mathcal{A}} \Big[(1-\alpha)\pi_a(\rho, p_{bt}) + \alpha\pi_a\big(\sigma(p_{bt}), p_{bt}\big)\Big], \tag{16}$$

reflecting the simultaneous Bertrand incentive (first term) and the within-period response incentive (second term). 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 set its punishment function equal to its static best response, which coincides with Assumption A2, and to choose $\rho = R_a(p_b^\dagger)$, so the second term corresponds to the Stackelberg follower incentive. The slower firm understands that, for the last α fraction of the period, firm a will have already reacted to its price. Under standard regularity conditions, if $\beta\gamma = 0$ and $\alpha \in (0, 1)$, the equilibrium price vector lies on firm a's static best-response function between the Bertrand outcome and (p_a^F, p_b^L), and prices exceed the Bertrand benchmark when prices are strategic complements.

Maximal Coercion The maximal coercion problem in (14) is qualitatively different from the two sequential benchmarks. Neither firm plays its static best response. Instead, firm a chooses the most profitable price pair that gives firm b at least its best deviation payoff. If firm b complies, it receives $\pi_b(p_b^\dagger, p_a^\dagger)$. If it deviates to $\hat{p}_b$, the algorithm immediately responds with $R_a(\hat{p}_b)$, so the right-hand side of (14) is the payoff firm b would obtain as the leader in a sequential-move pricing game. The speed-only Stackelberg pair (p_a^F, p_b^L) satisfies the incentive constraint with equality, but other price vectors can also be feasible. In equilibrium, firm a selects the feasible target pair that maximizes its own profit on or above this isoprofit constraint. The implementation of maximal coercion thus has a mechanism design flavor.

This mechanism differs from standard repeated-game punishment. In a standard repeated-game punishment model, the players respond to potential punishment in future periods. In contrast, the rival is induced to comply with the target prices based only on its current-period

payoffs. The algorithmic firm knows that it is playing against a myopic rival that cannot implement dynamic punishment strategies. The algorithmic firm is instead disciplined by commitment. Commitment allows the algorithmic firm to internalize the effect of reducing its price on its stream of future profits, when the rival updates its (static) best response.

With commitment only, firm a can commit to its own price but cannot make a rival's within-period deviation unprofitable. With speed only, firm a can react quickly but cannot credibly hold itself to a persistent high-price rule. With both commitment and speed, firm a can discipline the rival through rapid responses and discipline itself through commitment.

This logic translates to interior values of α and $\beta\gamma$, where coercion yields results that are not merely an interpolation between the two Stackelberg benchmarks. The objective function in (10) reflects a combination of simultaneous, leader, and follower incentives that interact with each other, allowing for the possibility of substantially higher prices. We explore this more in Section 4.

3.2 Collusive Prices and Profits

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

$$(p_a^C, p_b^C) \in \arg\max_{(p_a,p_b)} \ \pi_a(p_a, p_b) + \pi_b(p_b, p_a), \tag{17}$$

and we define each firm's *full collusion profits* as the profit it earns at these prices: $\pi_a(p_a^C, p_b^C)$ and $\pi_b(p_b^C, p_a^C)$.

The next result characterizes when the collusive prices can be implemented under the baseline algorithm. This benchmark does not, in general, describe the target chosen by the algorithmic firm, as it does not impose firm a's optimality.

Proposition 2 (Implementability of the Collusive Benchmark)**.** *Suppose Assumptions A1–A2 hold and there exists a joint-profit-maximizing price vector (p_a^C, p_b^C), defined in* (17)*, satisfying*

$$\pi_b(p_b^C, p_a^C) \geq \pi_b\big(\hat{p}_b, R_a(\hat{p}_b)\big) \qquad \textit{for all } \hat{p}_b \in P_b \setminus \{p_b^C\}. \tag{18}$$

Then there exists $\alpha^C \in [0,1]$ such that, for any $\alpha \in [\alpha^C, 1]$, the collusive price vector (p_a^C, p_b^C) is implementable as a target of the baseline algorithm.

Condition (18) requires firm b's profit at the collusive prices to weakly exceed its Stackelberg-leader profit. The condition can fail, for example, when demand conditions differ substantially across firms. If firm b has weak demand, then firm a's best-response to a deviation by firm b is not a sufficient threat to maintain collusion.

Proposition 2 characterizes when firm a could induce b to follow the joint-profit-maximizing prices. However, absent common ownership, transfers, or explicit agreement, firm a need not maximize joint profits. Instead, in coercive equilibrium, it solves (10), maximizing its own profit subject to firm b's incentive compatibility constraint.

The result helps to distinguish the separate roles of speed and commitment. Speed enters firm b's constraint directly and determines whether the collusive price vector is implementable, whereas speed and commitment together determine the target that firm a chooses in equilibrium. Thus, $\alpha \geq \alpha^C$ does not imply that firm a selects the collusive prices or earns its full collusion profits. Conversely, $\alpha < \alpha^C$ does not rule out firm a earning more than it would under full collusion at another target. When firm b's incentive constraint is slack at the collusive prices, firm a can often move to a more profitable target while preserving firm b's compliance. Maximal coercion can therefore give firm a more than its profit under full collusion even though total industry profit remains below the joint-profit maximum. Appendix Figure A-4 illustrates this distinction with an example demand system. We return to this discussion in Section 4.1.

3.3 Numerical Example with Linear Demand

We illustrate these benchmark cases with examples generated from the symmetric linear demand system

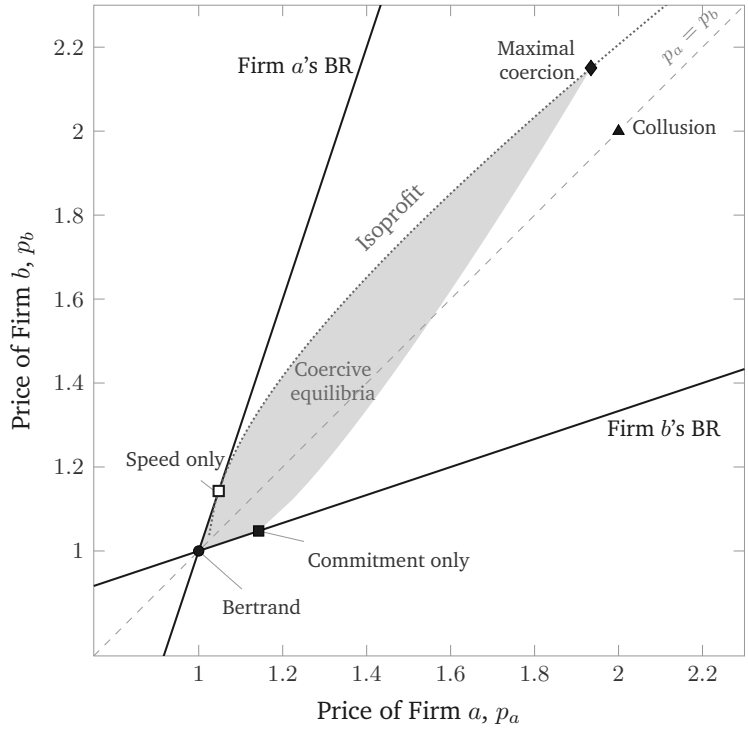
$$D_i(p_i, p_{-i}) = 1 - \left(\frac{1}{4} + \frac{d}{2}\right) p_i + \frac{d}{2} p_{-i}, \tag{19}$$

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). We normalize marginal costs to zero and restrict attention to the economically relevant price region in which $D_i(p_i, p_{-i}) \geq 0$ for each firm.

Figure 2 plots the four benchmark price pairs under the linear demand system in (19) with $d = 1$. With neither speed nor commitment, the unique outcome is the symmetric Bertrand–Nash equilibrium, $(p_a, p_b) = (1, 1)$. With full commitment but no speed, firm a acts as the leader in a Stackelberg pricing game and the equilibrium is $(p_a, p_b) = (1.14, 1.05)$. With maximal speed but no commitment, the roles are flipped: firm a acts as a Stackelberg follower and the equilibrium is $(p_a, p_b) = (1.05, 1.14)$. These one-feature benchmarks yield moderate price increases above Bertrand (5–14 percent).

With speed and commitment, prices can be significantly higher. The maximal-coercion case, with $(\alpha, \beta\gamma) = (1, 1)$, generates equilibrium prices $(p_a, p_b) = (1.93, 2.15)$. This is the most profitable point for firm a consistent with firm b's incentive constraint, which is displayed in the figure as an isoprofit curve (dotted line). The algorithmic firm prices below its rival, but both prices are far above the Bertrand benchmark. The quantity-weighted average price is 2.01. For comparison, joint profits are maximized at $(p_a, p_b) = (2, 2)$. Thus, average coercive prices slightly exceed collusive prices.

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 indicates the price for firm a and the vertical axis indicates 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 isoprofit curve shown indicates the points that provide firm b its sequential-leader profit. Demand follows the linear specification in equation (19) with $d = 1$.

The algorithmic firm implements the maximal-coercion outcome using an update rule that coincides with its static best response whenever firm b deviates from its target price:

$$\sigma(p_{bt}) = \begin{cases} 1.93 & \text{if } p_{bt} = 2.15, \\ (2 + p_{bt})/3 & \text{if } p_{bt} \neq 2.15. \end{cases} \tag{20}$$

If firm b sets $p_{bt} = 2.15$, its profit is $\pi_b(2.15, 1.93) = 0.76$. If it deviates, it faces firm a's 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 payoff. Thus the incentive compatibility constraint binds at the target pair.[17]

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

[17]Firm b's optimal deviation from the maximal-coercion equilibrium is to choose a price of 1.14, which induces firm a's algorithm to set a price of 1.05. These are the sequential leader and follower prices. Note that prices are rounded to two decimal places.

raises firm a's profits by about 61 percent. Moreover, the algorithmic firm earns strictly more 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 studies this attainable set more systematically.

4 Prices, Profits, and Welfare in Equilibrium

In this section, we consider prices, profits, and consumer surplus over a range of technology parameters. We address private incentives to adopt pricing technology in our context and provide extensions to harsher punishment rules, price matching, and oligopoly settings. The analysis shows that the economic mechanism is robust along several dimensions. Later, in Section 5, we relax the full-information assumption and consider rivals that employ naive learning rules.

4.1 Equilibrium Prices and Profits

We now study how coercive equilibrium outcomes vary with the algorithmic firm's speed and commitment. The previous section produced two striking properties of coercive equilibria:

Remark 1. *In a coercive equilibrium, the rival's price can exceed its joint-profit-maximizing price.*

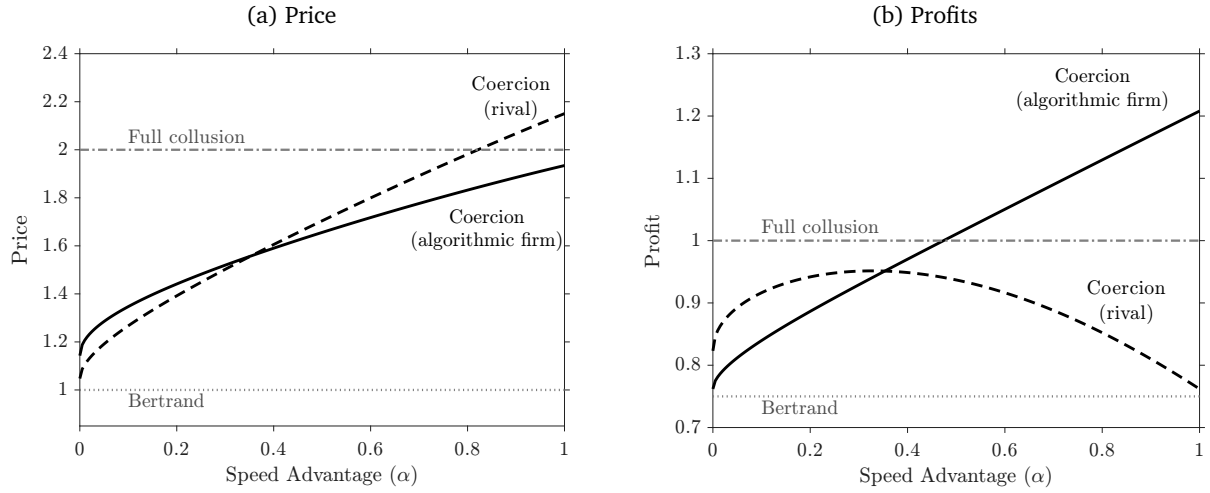
Remark 2. *In a coercive equilibrium, the algorithmic firm can earn more than its full collusion profits.*

Here, we examine how these properties arise with intermediate values of speed and commitment. Speed increases the effectiveness of punishment, while commitment reduces the algorithmic firm's temptation to cut prices for short-run gains. When combined, even at modest levels, they can reinforce each other and appreciably increase prices and profits.

We first provide illustrations using the linear demand system in equation (19) with $d = 1$, so that the benchmark cases in Section 3.3 provide a convenient comparison. We then characterize general properties of the equilibrium.

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

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



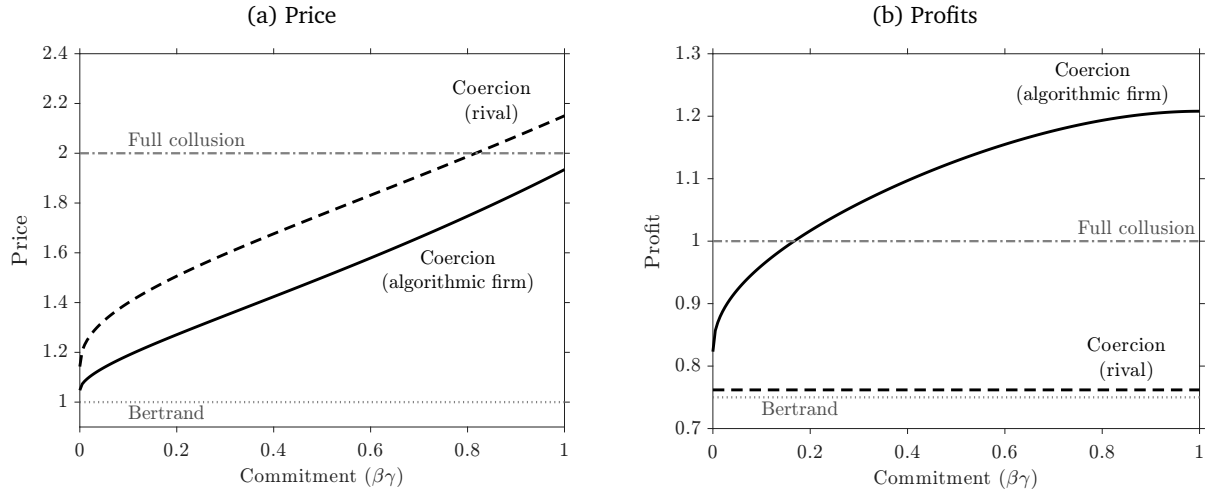
Notes: Panel (a) shows equilibrium prices and panel (b) shows equilibrium profits under Bertrand competition with simultaneous pricing, coercion with full commitment ($\beta\gamma = 1$), and joint-profit maximization. The figure reports outcomes for different values of the speed advantage α of the algorithmic firm. Demand follows the linear specification in equation (19) with $d = 1$.

commitment and α is sufficiently large (approximately $\alpha > 0.8$), the slower rival's equilibrium price lies strictly above the collusive level. 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 values of α above approximately 0.5, the algorithmic firm earns more than its full collusion profits, as seen in panel (b). In this region, 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 effective commitment $\beta\gamma$ varies. As commitment increases, both firms' prices rise above the Bertrand benchmark. With weak commitment, the algorithmic firm places substantial weight on the current-period gain from cutting its price, which leads to lower equilibrium prices. As $\beta\gamma$ increases, future profit becomes more important, allowing the algorithmic firm to commit itself to a higher price. The algorithmic firm earns more than its full collusion profit when $\beta\gamma$ exceeds approximately 0.2.

These figures illustrate how a speed advantage and commitment across periods have different effects on feasible profits. For the algorithmic firm, profits are monotonically increasing in speed and commitment. However, the profit implications for the rival differ. As shown in Figure 3, when the algorithmic firm's speed advantage is small, the rival earns greater profits than the algorithmic firm. Around $\alpha = 0.35$, the lines for firm a and firm b intersect. This reflects the point at which the leader-like effect of commitment and the follower-like effect of speed balance each other. By contrast, in Figure 4, stronger commitment increases profits, but only for the algorithmic firm; profits for the rival never exceed the sequential-move level. When

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 $\beta\gamma$. Demand follows the linear specification in equation (19) with $d = 1$.

$\alpha = 1$, the rival's payoff is pinned down by the immediate best-response punishment.[18]

For this example, speed and commitment reinforce one another in raising equilibrium prices. With commitment alone or a speed advantage alone, the algorithmic firm's ability to raise prices and profits is bounded by the corresponding sequential-move outcome. Without commitment, the algorithmic firm is tempted to best-respond to the rival's price for short-term gain. 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-5 and A-6. For example, the algorithmic firm earns more than its full collusion profits if either $\beta\gamma = 0.5$ or $\alpha = 0.5$, provided a sufficient amount of speed or commitment, respectively.

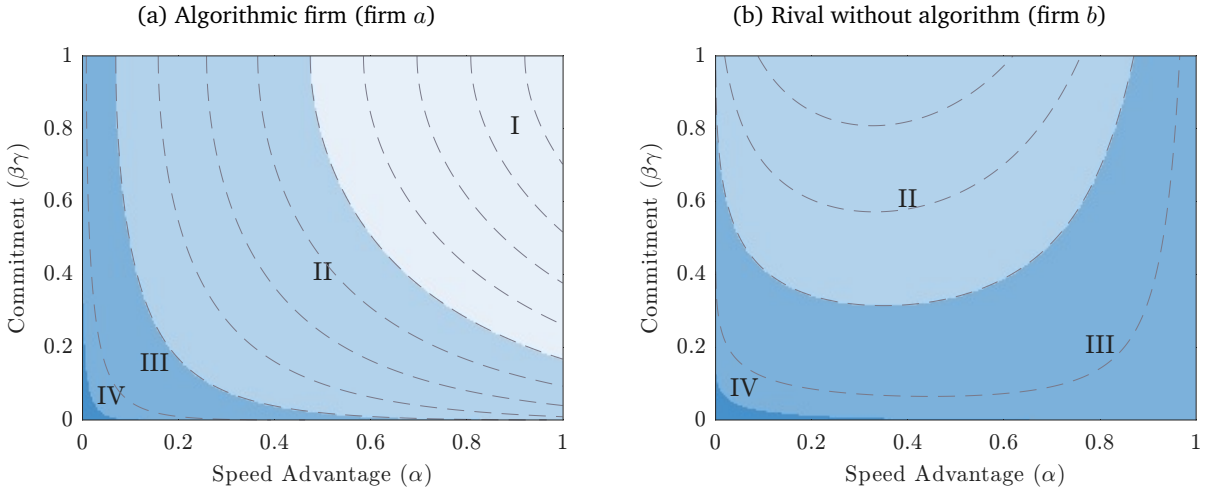
This technological complementarity can be formalized with a simple profit comparison. For a fixed value of the product-differentiation parameter d, let $\Pi_a(\alpha, \beta\gamma)$ denote firm a's equilibrium per-period profit and define

$$\Delta(\xi) \equiv \Pi_a(\xi, \xi) - \Pi_a(\xi, 0) - \Pi_a(0, \xi) + \Pi_a(0, 0), \tag{21}$$

for a technology level $\xi \in [0, 1]$, so that $\alpha = \beta\gamma = \xi$. Positive values of $\Delta(\xi)$ mean that

[18] Appendix Figure A-7, which shows prices and profits with logit demand, and Appendix Figure A-8, which explores different values of product differentiation, show that the qualitative forces of the mechanism are not driven by the linear-demand specification or the baseline value of product differentiation.

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 profit; 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 the Bertrand benchmark. 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. Demand follows the linear specification in equation (19) with $d = 1$.

combining speed and commitment raises profit by more than the sum of the gains from each technology in isolation.[19] In our linear demand example, $\Delta(\xi) > 0$ for every $\xi \in (0, 1]$ and $d > 0$. Appendix A.3 reports this measure, along with a version relative to the value without commitment, for different values of ξ and d.

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 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 the profit implications of speed and commitment. We show that, when the

[19] Equivalently, $\Delta(\xi) > 0$ can be written as $\Pi_a(\xi,\xi) - \Pi_a(0,0) > \big(\Pi_a(\xi,0) - \Pi_a(0,0)\big) + \big(\Pi_a(0,\xi) - \Pi_a(0,0)\big)$, or either marginal-value comparison: $\Pi_a(\xi,\xi) - \Pi_a(0,\xi) > \Pi_a(\xi,0) - \Pi_a(0,0)$ and $\Pi_a(\xi,\xi) - \Pi_a(\xi,0) > \Pi_a(0,\xi) - \Pi_a(0,0)$.

algorithm acts as a punishment, firm a can secure weakly higher profits as its speed advantage increases. In addition, firm a's per-period profit is weakly increasing in effective commitment.

Proposition 3 (Profit Monotonicity in Speed)**.** *Fix effective commitment $\beta\gamma$. Consider the set of equilibrium target pairs and the optimal deviations from these target pairs by firm b for all values of α. If each optimal deviation results in an algorithm update that weakly lowers firm b's payoff, then firm a's securable profits are weakly increasing in α.*

Proposition 4 (Profit Monotonicity in Commitment)**.** *Fix the speed advantage α and the current rival price p_{bt}. The algorithmic firm's per-period profit at its optimal target pair is weakly increasing in effective commitment $\beta\gamma$.*

Intuitively, a higher α can strengthen the effectiveness of punishment following a deviation. A more effective punishment relaxes firm b's incentive constraint and allows firm a to choose a more profitable target pair. The additional condition stated in the proposition requires the algorithm's update to reduce firm b's payoff relative to maintaining firm a's target price. Instead, if the algorithm were to respond by raising prices, a faster reaction would make deviations more profitable and tighten the incentive constraint.

The result for commitment does not require the additional condition that an update acts as a punishment. The mechanism is distinct: a higher level of commitment $\beta\gamma$ reduces the temptation to cut its own price in the current period because future profits under the algorithm become more important.

Taken together, the results indicate the distinct roles of speed and commitment. Speed can relax firm b's incentive constraint and increase the profit firm a can attain, whereas commitment determines the target prices chosen within the incentive-compatible set. Proposition 3 characterizes the feasible profits that can be secured when α increases, relative to equilibrium target prices. Proposition 4 holds fixed the current rival price at the reprogramming date, isolating the effect of commitment on firm a's target-price problem. In equilibrium, the current rival price also responds to changes in the technology parameters. The algorithmic firm's equilibrium per-period profit is increasing in both speed and commitment under linear demand, which we present as Lemma 3 in the Appendix.

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

Proposition 5 (Supracollusive Profits)**.** *Let (p_a^C, p_b^C) denote an interior joint-profit-maximizing (collusive) price vector, and let (p_a^F, p_b^L) denote the one-shot Stackelberg outcome in which firm b moves first and firm a best responds. If $\pi_b(p_b^C, p_a^C) > \pi_b(p_b^L, p_a^F)$, then there exist thresholds $\overline{\alpha} \in (0,1)$ and $\overline{\beta\gamma} \in (0,1)$ such that for all $\alpha > \overline{\alpha}$ and $\beta\gamma > \overline{\beta\gamma}$, 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 for sufficiently high α the collusive price vector satisfies firm b's current-period 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 while still deterring deviations. In this sense, coercion allows the algorithmic firm to use the threat of rapid undercutting to extract surplus from the rival.

Following the discussion in Section 3.2, the region where $(\alpha, \beta\gamma)$ yields supracollusive profit for firm a need not correspond to the region in which the full collusion price vector is incentive compatible for firm b. Appendix Figure A-4 illustrates the distinction. Collusive implementability depends only on speed; for our example demand system, collusion is incentive compatible for firm b when $\alpha > 0.573$. Whether or not equilibrium profits for firm a exceed full collusion profits depends on both speed and commitment. For some values of α below this threshold, it is possible for firm a to earn more than its full collusion profits. However, when $\alpha > 0.573$ but commitment is low, the algorithmic firm may earn less than its full collusion profits. In such cases, firm a sacrifices long-run gains to chase short-run profits.

The above propositions apply under our general assumptions, but comparative statics for prices and welfare require additional restrictions about demand. We now impose the linear-demand specification used in the numerical example from Section 3.3.

Assumption D1 (Linear Demand and Zero Marginal Cost)**.** *Demand follows the linear specification in equation* (19)*, and both firms have zero marginal cost.*

Under this specification, we can also establish monotonicity of equilibrium target prices within the class established by Assumption A1, Assumption A2, and Appendix Assumption R1. Demand systems with different curvature properties may not preserve monotonicity in prices.

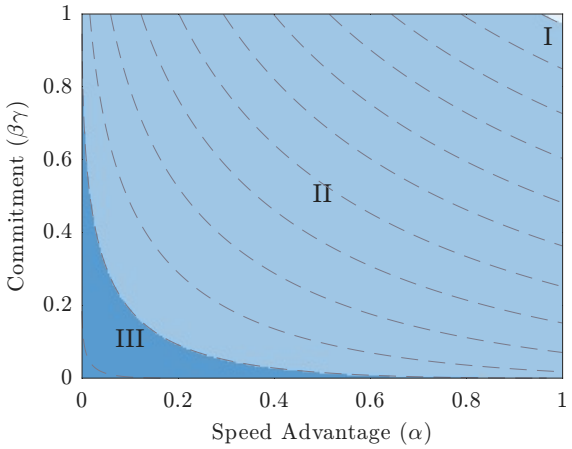
Proposition 6 (Price Monotonicity in Speed and Commitment)**.** *Under Assumption D1, both firms' equilibrium target prices $(p_a^\dagger, p_b^\dagger)$ are strictly increasing in the speed advantage α for each fixed $\beta\gamma$ and strictly increasing in the effective commitment parameter $\beta\gamma$ for each fixed α.*

Key conditions for this to hold are that the best 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 firm's best response than facing the target price: $\pi_b(\hat{p}_b, R_a(\hat{p}_b)) < \pi_b(\hat{p}_b, p_a^\dagger)$. Assumption D1 is a convenient sufficient assumption under which the two conditions can be verified in closed form. These conditions state that a faster algorithm reduces the rival's payoff from deviating by reacting more quickly. This relaxes the rival's incentive compatibility constraint and enables the algorithmic firm to sustain higher target prices.

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

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 (19) with $d = 1$.

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 bottom-left corner is the Bertrand outcome, while the small upper-right area identifies the parameter combinations for which consumers fare worse than under collusion. More generally, 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. Therefore, total welfare is also lower. We summarize these observations as follows.

Remark 3. *Consumer surplus and total welfare can be lower in a 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 rival is coerced into setting a price that can exceed the collusive level, while the algorithmic firm prices slightly below. Since the products are differentiated, these prices can 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 with speed and commitment. One sufficient condition is that demand is linear. Proposition 6 establishes that both target prices are increasing in speed and commitment. Since consumer surplus is decreasing in prices, increasing speed or commitment for the algorithmic firm leads to a monotonic decline in consumer surplus.

Proposition 7 (Consumer Surplus Monotonicity)**.** *Under Assumption D1, equilibrium consumer surplus is strictly decreasing in α and in $\beta\gamma$.*

The above analysis shows that the combination of faster pricing and (imperfect) commitment can have first-order consequences for consumers. In the next subsection, we show that this qualitative finding extends to alternative environments.

4.3 Extensions and Robustness

This subsection presents extensions that explore the robustness of the coercive mechanism. First, we consider private incentives that lead to the underlying asymmetric technology. Second, we examine harsher punishments than the best-response punishment of Assumption A2. Third, we explore price matching, a simple alternative pricing rule. Finally, we examine an oligopoly setting with one algorithmic firm and several slow rivals. Sections 5 and 6 present additional extensions.

Private Incentives and Endogenous Asymmetry

The comparative statics in Section 4.1 speak to private incentives to acquire pricing technology. In the linear-demand environment, Lemma 3 implies that, holding firm b's technology fixed, firm a's equilibrium profit is increasing in its speed advantage α and in its effective commitment $\beta\gamma$. Thus, absent adoption costs, it is optimal for firm a to choose the greatest available speed and commitment. With adoption costs, a firm invests whenever the resulting increase in equilibrium profit exceeds the incremental cost.

The slower firm's incentives are more subtle. Consider a case in which firm a has obtained a speed advantage, and firm b can invest in higher-frequency pricing to reduce this advantage. Firm b can match but not exceed firm a's speed, as would be the case if firm a had the fastest pricing technology available. Reducing firm a's speed advantage weakens its ability to discipline deviations, but it also makes price competition more intense. In the linear-demand example, holding firm a's commitment fixed at $\beta\gamma = 1$, firm b's profit is maximized when firm a has a modest speed advantage. This asymmetry softens competition and allows both firms to share in the additional profits from higher prices.

Remark 4. *In a coercive equilibrium, the slower firm's profit can be maximized at a strictly positive speed advantage for firm a.*

Appendix A.4 formalizes this calculation and embeds it in a simple adoption game.[20]

The analysis above assumes that firm b remains the slower, non-committed rival. If firm a is not at the frontier, so that firm b can obtain the technological lead by leapfrogging its rival, the additional profit earned by the faster firm creates the possibility of an "arms race" in pricing technology, analogous to speed races in high-frequency trading (Biais et al., 2015; Budish et al., 2015; Aquilina et al., 2022). In addition, if firm b instead adopts a fast and persistent pricing technology, the environment becomes one of competition among algorithms rather than one in which a committed fast firm disciplines a slow rival. In that case, firm b can also write rules that respond to firm a's prices and can commit to those rules over time. Brown and MacKay (2023) study one such environment and show that simple linear rules that condition on rivals' prices can support supracompetitive prices, including collusive outcomes.

Harsher Punishments

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 collusion literature. It is, however, not the only possibility. With a harsher punishment, firm a can further reduce firm b's payoff from deviating and increase its own profits.

Consider the ordered class of punishments

$$\sigma_\eta(p_b) \equiv (1-\eta)\, R_a(p_b) + \eta\, \underline{p}_a, \qquad p_b \neq p_b^\dagger, \quad \eta \in [0,1], \tag{22}$$

with $\sigma_\eta(p_b^\dagger) = p_a^\dagger$ at the target, where $\underline{p}_a$ denotes the lowest feasible price for firm a, which we assume is equal to its marginal cost. Higher values of η correspond to harsher punishments: $\eta = 0$ recovers the best-response punishment of Assumption A2, and $\eta = 1$ is the lowest-price

[20] The analysis isolates the speed margin by holding firm a's effective commitment fixed. In calendar time, a change in firm b's pricing frequency can also change the mapping from time to model periods. Thus, investing in speed both reduces firm a's speed advantage and increases the per-period persistence associated with the same calendar-time commitment.

punishment. Because the products are substitutes, the lowest-price punishment gives firm b the lowest payoff.

Proposition 8 (Punishment Intensity)**.** *For the class of punishment rules $\sigma_\eta(\cdot)$ above, firm b's best deviation payoff is weakly decreasing in the punishment intensity η. Hence any target pair implementable under a milder punishment is also implementable under any harsher punishment, and the lowest-price punishment provides an upper bound on securable profits.*

Proposition 8 indicates that marginal-cost punishment provides the largest securable profits within the considered class. For this reason, we view best-response punishment as conservative. Appendix Figure A-9 illustrates equilibrium prices and profits for the harshest punishment case for our linear-demand example ($\eta = 1$ and $\underline{p}_a = 0$). The qualitative relationships with the speed advantage α and commitment $\beta\gamma$ are similar to those under best-response punishment, but prices and the algorithmic firm's profit are even higher. With marginal-cost punishment, the algorithmic firm can induce firm b to earn less than its Bertrand profits.

Price Matching

We next consider a qualitatively different punishment in which the algorithmic firm employs price matching, so that $\sigma(p_b) = p_b$ following a deviation from the target price. Price-matching rules or guarantees are observed in practice and may soften competition (see, for example, Salop, 1986). Under price matching, equilibrium prices and profits increase with the speed advantage and commitment for our symmetric linear-demand example. However, when firm a can react instantaneously, the coercive equilibrium coincides with the fully collusive outcome, regardless of the level of commitment (Appendix Figure A-10). 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, our baseline best-response punishment delivers larger gains to the algorithmic firm when commitment is strong, since it allows firm a to earn more than its full collusion profits. By contrast, with weak commitment, price matching can yield higher profits than best-response punishment.

One Algorithmic Firm in an N-Firm Oligopoly

We next extend the analysis of coercion to an oligopoly setting. Consider an N-firm market with one algorithmic firm, denoted a, and $N-1$ "slow" rivals that do not use algorithms. Let $\mathbf{r} = (r_1, \ldots, r_{N-1})$ denote the vector of slow rivals' prices, and let $\mathbf{r}_{-j}$ denote the vector

excluding rival j's price. If firm a sets target prices $(p_a^\dagger, \mathbf{r}^\dagger)$, slow rival j's incentive constraint is

$$\pi_j(p_a^\dagger, r_j^\dagger, \mathbf{r}_{-j}^\dagger) \geq \max_{\hat{r}_j} \left\{ (1-\alpha)\pi_j(p_a^\dagger, \hat{r}_j, \mathbf{r}_{-j}^\dagger) + \alpha\pi_j\Big(R_a(\hat{r}_j, \mathbf{r}_{-j}^\dagger), \hat{r}_j, \mathbf{r}_{-j}^\dagger\Big)\right\}. \tag{23}$$

The left-hand side is rival j's payoff from charging its target price. On the right-hand side, the first term is its payoff before firm a reacts and the second is its payoff after firm a applies the punishment. Firm j considers these deviations while holding the prices of other slow rivals fixed at $\mathbf{r}_{-j}^\dagger$. Under marginal-cost punishment, the second term on the right-hand side is evaluated with the algorithmic firm's price set equal to marginal cost instead of $R_a(\hat{r}_j, \mathbf{r}_{-j}^\dagger)$. Appendix A.5 characterizes the full general oligopoly problem.

While similar to the duopoly incentive constraint, equation (23) shows how additional competitors can weaken coercion. When a slow rival deviates, the other slow rivals remain at their target prices. Thus a deviating firm can steal demand not only from the algorithmic firm, whose price responds quickly, but also from other slow firms, whose prices do not. This increases the deviation payoff for the slow rivals, leading the algorithmic firm to set target prices that are more favorable to its rivals in equilibrium.

Nevertheless, coercion can still yield substantial equilibrium impacts in an oligopoly setting. For transparent comparisons, we focus on maximal coercion and the symmetric linear-demand system

$$D_i(p_i, p_{-i}) = 1 - \left(\frac{1}{4} + \frac{d}{2}\right) p_i + \frac{d}{2}\overline{p}_{-i}, \tag{24}$$

where $\overline{p}_{-i}$ is the average price charged by firm i's rivals. This specification nests the duopoly example and preserves the symmetric benchmarks: the Bertrand price is $p^B(d) = 2/(1+d)$ and the joint-profit-maximizing price is $p^C = 2$.

Under this demand system, Appendix Figure A-3 shows that prices and profits can still be substantially above their Bertrand levels even with several slow rivals. With ten rivals, average prices are 18 percent higher than Bertrand prices under best-response punishment and 43 percent higher under marginal-cost punishment.

We formalize the effects of maximal coercion in oligopoly with a comparison to a cartel benchmark, using marginal-cost punishment. We focus on cartel implementability and firm a's profit because they admit closed-form comparisons. For $K \in \{2, \ldots, N\}$, consider the static benchmark in which firm a and $K-1$ rivals form a cartel and the remaining firms price noncooperatively.

Proposition 9 (Implementability of K-firm Cartels)**.** *Maintain the linear demand system in equation* (24)*, assume marginal costs are zero, and consider maximal coercion ($\alpha = 1$, $\beta\gamma \to 1$) with marginal-cost punishment.*

(i) For $K \in \{2, \ldots, N\}$, the K-firm cartel price vector is implementable if and only if $K^2 \leq 2(1+2d)(N-1)/d$. Therefore, a $K = 2$ cartel price vector is implementable for every N and

d.

(ii) Let $K^*(N,d)$ denote the largest K for which the K-firm cartel price vector is implementable. For $N \geq 3$, $K^*(N,d) = N$ if and only if $0 < d \leq 2(N-1)/(N-2)^2$. Thus, for every N, the grand cartel is implementable when products are sufficiently differentiated.

(iii) Firm a's maximal-coercion profit is at least its profit as a member of a $K^*(N,d)$-firm cartel. It is strictly greater than its profit as a member of a two-firm cartel and, when the inequality in part (ii) is strict, strictly greater than its profit as a member of the grand cartel.

Part (i) formalizes the effects of market structure and product differentiation on implementability. The left-hand side increases with the cartel size, whereas the right-hand side increases with product differentiation. A larger cartel is harder to enforce. With more firms setting higher prices, deviations are more attractive. Product differentiation works in the opposite direction: when slow rivals are sufficiently differentiated, a deviating rival steals relatively little from the other slow firms, so the algorithmic firm's rapid punishment remains effective. Thus, cartel implementability is determined by the interaction between the number of slow rivals and the strength of substitution among them. Note that a two-firm cartel is always implementable: a single algorithmic firm can induce at least one rival to charge the corresponding cartel price.

Part (ii) characterizes when the algorithmic firm could induce all firms to set collusive prices. For a fixed N, the grand cartel is implementable with a low value of d, indicating that products are sufficiently differentiated. The condition also provides comparative statics. For fixed d, $K^*(N,d)/N$ declines with N (up to integer rounding). For example, when $d = 1$, $K^*(N,d)$ equals 6 for $N = 7$, 7 for $N = 10$, and 10 for $N = 20$. For fixed N, $K^*(N,d)$ is weakly decreasing in d. When $d = 0.2$, the grand cartel is implementable for every $N \leq 12$, and $K^*(N,d) = 16$ for $N = 20$.

Part (iii) distinguishes implementability from the algorithmic firm's equilibrium choice. Any implementable cartel vector is feasible in the maximal-coercion problem, so firm a can always earn at least the associated cartel-participant profit. When the incentive constraints hold with slack, firm a can move to a nearby target that raises the slow rivals' prices and increases its own profit. This is why maximal coercion gives firm a strictly more than the two-firm cartel benchmark and, when the grand-cartel condition holds strictly, more than its grand-cartel profit.[21]

[21] The cartel condition is sufficient, but not necessary, for supracollusive profit. Firm a may earn more than its grand-cartel profit even when the grand-cartel price vector is not itself implementable. With $d = 1$, for example, the grand cartel is implementable with at most three slow rivals ($N \leq 4$), whereas firm a earns supracollusive profit through $N = 9$, as illustrated in Appendix Figure A-3, panel (d). This discrepancy further illustrates a difference between collusion and coercion in terms of the incentives for the adopting firm.

5 Learning by a Naive Rival

The analysis so far assumes that the rival understands the algorithmic firm's rule and best responds to it period by period. This section demonstrates how coercion can arise when this informational assumption is relaxed. In some markets, an algorithmic firm may face less sophisticated rivals that are not fully aware of the algorithm or the structure of demand. For example, a manager may instead run pricing experiments and simply adjust prices in the direction that appears to raise profits.

To capture this, here we treat firm b as naive in the sense that it observes only its own prices and realized profits and chooses future prices as a function of that history. First, we demonstrate how an algorithmic firm can lead a naive rival to coercive prices by using a linear pricing rule. Second, we show that convergence is fast in simulations, making the mechanism relevant even when firms experiment over a small number of pricing periods.

5.1 Naive Learning Rules

We assume that firm b observes its own prices and realized profits, but nothing else about the economic environment. We consider learning rules in which the firm uses this information to search for a profit-maximizing price. Let $H_t^b \equiv \left\{(p_{b,\ell}, \Pi_{b,\ell})\right\}_{\ell=0}^{t}$ denote firm b's history through period t, where $p_{b,\ell}$ is its posted price and $\Pi_{b,\ell}$ is its realized profit. Let $r_{b,t}$ denote the firm's current reference price. This reference price is simply the point around which the learning rule updates or experiments and need not be optimal. The firm chooses its next reference price according to

$$r_{b,t+1} = L_t(H_t^b) = L_t\big(\{p_{b,\ell}\}_{\ell=0}^{t}, \{\Pi_{b,\ell}\}_{\ell=0}^{t}\big)\,, \tag{25}$$

where L_t is fixed ex ante and does not adjust to the demand system or to firm a's algorithm. These rules require no knowledge of demand, no observation of the algorithmic firm's pricing rule, and no dynamic reasoning.[22]

Many learning approaches in this class can be written in the form

$$r_{b,t+1} = r_{b,t} + \lambda_t(H_t^b)\widehat{g}_{b,t}(H_t^b), \tag{26}$$

where $\lambda_t > 0$ is a step size and $\widehat{g}_{b,t}$ provides the sign and magnitude of an adjustment signal, such as the gradient of the profit function. For some methods, the reference price is the next selected price, $p_{b,t} = r_{b,t}$. Other methods might set experimental prices around $r_{b,t}$.

As a motivating example, we focus on finite-difference gradient learning. The slow firm

[22]This follows the learning-in-games perspective that stable behavior may emerge through adaptive experimentation and payoff feedback rather than full knowledge or strategic behavior. See Fudenberg and Levine (2016). Related work studies richer model-free procedures, such as Q-learning (Calvano et al., 2020), that learn history-dependent reward–punishment strategies. Here, by contrast, firm b simply attempts to maximize single-period profits rather than learn a state-contingent dynamic strategy.

starts with two distinct initial prices $p_{b,0} = r_{b,0}$ and $p_{b,1} = r_{b,1}$, and uses the corresponding change in realized profits to estimate the slope of its perceived profit function. For $t \geq 1$, the firm estimates

$$\widehat{g}_{b,t} = \frac{\Pi_{b,t} - \Pi_{b,t-1}}{p_{b,t} - p_{b,t-1}}, \tag{27}$$

and updates following

$$p_{b,t+1} = r_{b,t+1} = r_{b,t} + \lambda \widehat{g}_{b,t}, \tag{28}$$

with $\lambda = \lambda_t > 0$. Heuristically, the rule continues in the direction of the last profitable price change and reverses after an unprofitable change. If two distinct prices yield the same realized profit, we assume that firm b uses their midpoint as its next reference price. Throughout, the step size λ is small enough to keep the learning sequence stable. In simultaneous Bertrand settings, related gradient-learning rules converge to the Bertrand–Nash equilibrium provided λ is small enough (Anufriev et al., 2013).

Other learning approaches use the limited history in different ways. A quasi-Newton rule uses the three most recent price-profit observations. Specifically, it compares the profit slopes over the two latest price intervals and uses their change to determine the next price adjustment. Adaptive directional price search uses only the sign of a profit improvement. This approach continues in a profitable direction and reverses direction after an unprofitable experiment. Paired A/B testing posts prices just below and just above a reference price in successive periods. Appendix A.7 gives the precise update equations for these methods.

5.2 Linear Punishment Strategies

The presence of learning on the rival side also changes the algorithmic firm's problem. With learning, the algorithmic firm must consider whether its rival can reach the intended target. The baseline best-response punishment function, or any punishment function with a discontinuous trigger, can introduce a discontinuity in the rival's perceived profit function that may hinder convergence.

We restrict firm a to linear punishment strategies, which can make the induced optimization problem for firm b well behaved. In contrast to the discontinuous trigger rule in Assumption A2, 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

$$\sigma(p_b) = \max\{p_a^\dagger - \phi(p_b^\dagger - p_b), 0\}. \tag{29}$$

The punishment depends on the deviation $p_b - p_b^\dagger$ from the target price, and the intensity of the response is governed by $\phi > 0$. We focus on linear rules that pass through the origin. This implies $\phi = p_a^\dagger / p_b^\dagger$ with an update rule $\sigma(p_b) = \phi p_b$ for nonnegative prices. By construction, $\sigma(p_b^\dagger) = p_a^\dagger$.

The linear rule provides a continuous counterpart to the best-response and marginal-cost trigger rules discussed earlier: when firm b cuts its price, the algorithm lowers firm a's price proportionally. Algorithms that are linear functions of rivals' prices—including rules based on an average of competitor prices or rules that undercut observed prices by a fixed percentage—are common in online marketplaces (Chen et al., 2016; Musolff, 2024).

We maintain full effective commitment ($\beta\gamma = 1$) while the rival learns: firm a chooses the rule once to optimize steady-state profits and does not re-optimize it in response to the realized learning path. Conditional on this fixed rule, firm b's induced per-period profit as a function of its own price, or *perceived profit function*, is

$$\widetilde{\pi}_b(p_b) \equiv (1-\alpha)\pi_b(p_b, p_a^\dagger) + \alpha\pi_b(p_b, \phi p_b). \tag{30}$$

The linear pricing rule helps ensure that the naive rival's perceived profit function is well behaved. We formalize this property in the following subsection.

The algorithmic firm optimally chooses a target vector subject to the requirement that $p_b^\dagger$ maximize firm b's perceived profit. The target is therefore characterized by two conditions: the slower rival's perceived marginal profit is zero at $p_b^\dagger$, and the algorithmic firm optimally chooses $p_a^\dagger$ subject to that condition. Appendix A.6 gives the corresponding first-order conditions. Appendix Figure A-11 illustrates comparative statics for the linear-rule target. Linear algorithms preserve the economic logic of the baseline coercive equilibrium. Prices for both firms lie above the Bertrand benchmark and increase with the algorithmic firm's speed advantage α, and the algorithmic firm's profits can exceed its full collusion profits. However, restricting the algorithm to a linear rule supports less extreme target prices and somewhat lower profit for firm a than the discontinuous best-response trigger rule.

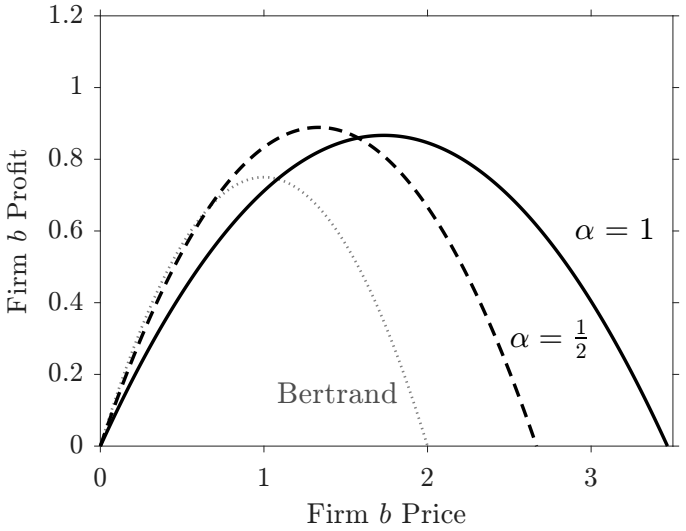
5.3 Convergence Dynamics

A linear pricing rule can make the mapping from firm b's own price to its perceived profit smooth and single-peaked. With the curvature condition below, firm b faces a one-dimensional choice problem with a strictly concave objective. Under linear demand and the optimal linear rule described above, the perceived profit function is exactly quadratic. This structure facilitates learning: even a simple adjustment rule that ignores the opponent's strategy and treats the environment as static can push prices toward the coercive target. This subsection formalizes this observation and illustrates it with simulations.

Lemma 2 (Single-Peaked Perceived Profit Function)**.** *Suppose that the algorithmic firm imposes the linear pricing rule above and that, for all relevant p_b, the curvature condition below holds:*

$$(1-\alpha)\,\frac{\partial^2\pi_b(p_b,\, p_a^\dagger)}{\partial p_b^2} + \alpha\left[\frac{\partial^2\pi_b(p_b,\, \phi p_b)}{\partial p_b^2} + 2\phi\,\frac{\partial^2\pi_b(p_b,\, \phi p_b)}{\partial p_b\,\partial p_a} + \phi^2\,\frac{\partial^2\pi_b(p_b,\, \phi p_b)}{\partial p_a^2}\right] < 0. \tag{31}$$

Figure 7: 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 simultaneous-pricing benchmark, in which the rival's profit is maximized at 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 its maximizer shifts to the right as α increases. Demand follows the linear specification in equation (19) with $d = 1$.

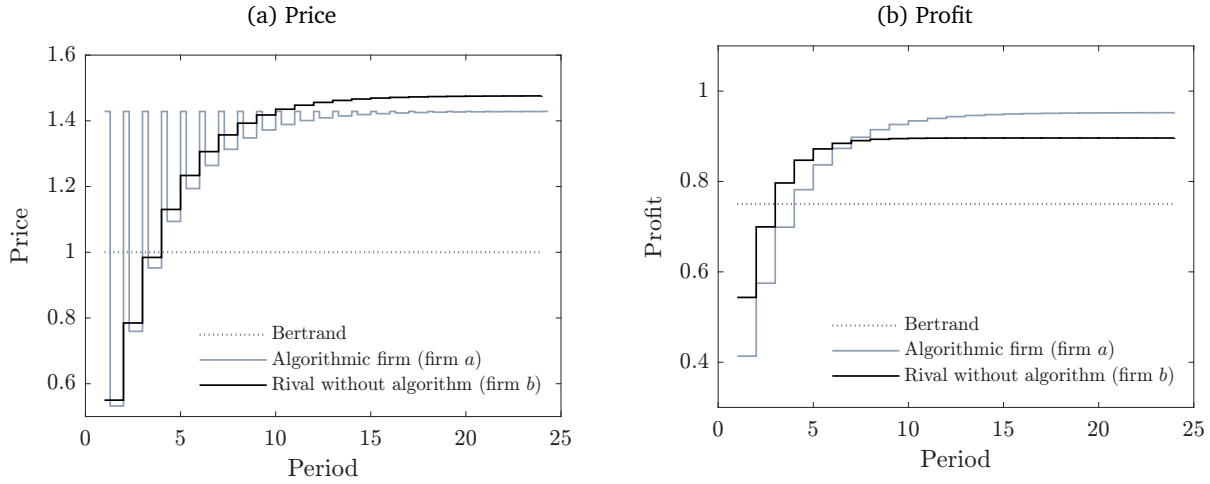
Then $\widetilde{\pi}_b$ is strictly concave. If $p_b^\dagger$ satisfies $d\widetilde{\pi}_b/dp_b = 0$, it is the unique maximizer of $\widetilde{\pi}_b$. Moreover, when demand is linear as in Assumption D1, $\widetilde{\pi}_b$ is a strictly concave quadratic.

The curvature condition ensures that the direct curvature of firm b's profit dominates the strategic feedback generated by firm a's response. The proof verifies the condition under the optimal linear rule for the linear demand system and gives the resulting quadratic explicitly.

When firm a uses a linear pricing rule and has a speed advantage, it effectively transforms firm b's problem by rotating the rival's perceived residual demand curve, making it appear more inelastic and diminishing the incentive to undercut. Firm b then perceives its profit-maximizing price to be higher. The linear rule helps ensure that firm b has a well-behaved objective, so that a solution can be found with simple rules.

Figure 7 illustrates the perceived profit function induced for the naive rival in the linear-demand example. Under simultaneous pricing, when firm a sets the Bertrand price $p_a = 1$, firm b's profit is 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, 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 firm b's perspective, it is optimal to raise price above the competitive level; this is entirely induced by the algorithmic firm's rapid response.

Figure 8: Simulated Learning with Linear Pricing Rule



Notes: The figure shows simulated price and profit paths when the naive firm uses finite-difference gradient learning. Firm b posts $p_{b,t} = r_{b,t}$ and updates the reference price according to $r_{b,t+1} = r_{b,t} + \lambda(\Pi_{b,t} - \Pi_{b,t-1})/(r_{b,t} - r_{b,t-1})$ with step size $\lambda = 0.3$. Firm a uses the linear pricing rule in Section 5.2 with speed advantage $\alpha = 0.7$. The dashed lines show prices and profits under the Bertrand benchmark. The simulation assumes $d = 1$ under the linear demand system in equation (19).

Because perceived profit is strictly concave in firm b's own price, the coercive target $p_b^\dagger$ is its unique maximizer. As an illustrative example, we formalize how finite-difference gradient learning converges to the desired target when λ is sufficiently small.

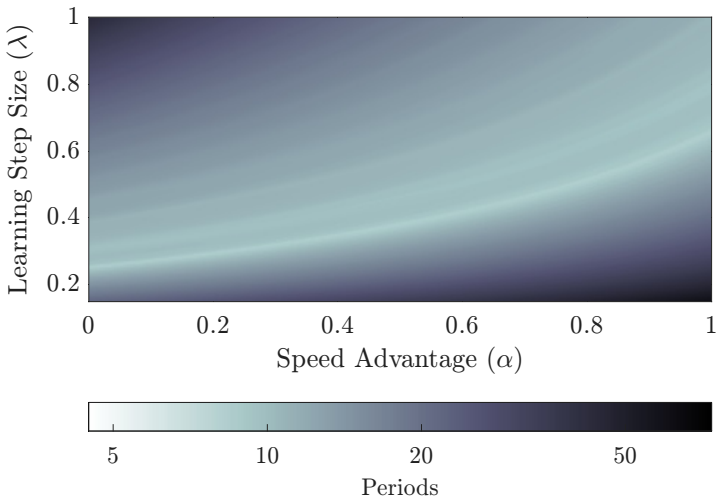
Proposition 10 (Convergence with Learning)**.** *Under Assumption D1, suppose firm a implements the optimal linear pricing rule and firm b follows the gradient learning rule in Section 5.1. If λ is sufficiently small ($0 < \lambda < \frac{4}{1+2d}$), then prices converge to the target prices $(p_a^\dagger, p_b^\dagger)$ from any two distinct initial prices in $[0, 2p_b^\dagger]$.*

The upper bound on λ provides a specific threshold to guarantee convergence, while the interval for initial prices, $[0, 2p_b^\dagger]$, is maintained to ensure that reference prices do not go negative.

Convergence can be fast, as we illustrate in Figure 8. For the simulation, firm a implements the linear pricing rule and has a moderate speed advantage, $\alpha = 0.7$. Firm b starts from two prices below the Bertrand level. At each step, it uses the two most recent price-profit observations to form the finite-difference estimate in equation (27) and then updates according to equation (28) with $\lambda = 0.3$. Panel (a) shows that firm b steadily raises its price while firm a responds through the linear pricing rule. Panel (b) shows the corresponding profit paths. As long as 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 estimated gradient remains positive along the plotted path as price rises toward the coercive target $p_b^\dagger$. In the limit, the system reaches a fixed point at the target prices.[23]

[23] Appendix Figure A-12 shows convergence from multiple starting prices and compares the coercive linear rule

Figure 9: Convergence Speed by Pricing Speed and Learning Step Size



Notes: The figure shows the number of periods until convergence for different simulation parameters when the naive firm uses 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 equally spaced initial reference prices $r_{b,0} \in [0, 2]$, set $r_{b,1} = r_{b,0} + 0.05$, and report the average number of periods to convergence. Convergence is defined as remaining within 0.1% of the steady-state target. Darker shading indicates more periods. The simulations assume that firm a uses the linear pricing rule in Section 5.2 and that $d = 1$ in the linear demand system in equation (19).

Across different initial prices, convergence occurs quickly, though the number of iterations depends on the algorithm's speed advantage and the learning step size λ. Figure 9 shows the average number of periods required for convergence across values of α and λ, using 1,000 simulated price paths for each pair of parameters. Consistent with Proposition 10, convergence is always achieved, and the number of iterations remains modest at every combination.[24] For every value of α, a learning parameter λ can be chosen to produce convergence in fewer than 10 periods on average.[25]

Convergence is not specific to finite-difference gradient learning. Appendix A.7 also reports simulations with paired A/B price experiments, quasi-Newton learning, and adaptive directional search in the same deterministic calibration. Using a common convergence criterion, all four methods converge from every initial price in the simulation grid. Mean convergence times are 4.0 periods for quasi-Newton learning, 20.2 for finite-difference gradient learning, 25.4 for directional search, and 46.9 for paired A/B experiments (Appendix Table A-3). Thus, for this example, convergence to the coercive target is robust across the simulated learning rules,

with a simultaneous-pricing benchmark.

[24]We identify convergence as the first of ten successive iterations in which prices are within 0.1% of the target. Given the nature of gradient learning, prices will not drift outside this tolerance in future periods. Appendix Table A-5 reports convergence statistics under more and less stringent criteria.

[25]By contrast, more complex reinforcement-learning algorithms that accommodate dynamic strategies can take thousands of iterations to approach a stable outcome and may require offline training before deployment (Calvano et al., 2020).

although convergence speed varies across methods.

The simulations illustrate how an algorithmic firm can reshape the perceived profit of its rival into a simple, well-behaved function with a peak at the coercive target. With naive learning, firm b does not observe firm a's price or pricing rule, yet it still quickly learns to price at the coercive outcome. More general pricing algorithms could reshape the rival's perceived profit function more flexibly. A nonlinear rule could potentially steer naive learning toward still higher prices or create a globally concave perceived objective under more general demand systems.

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. A platform can influence a seller's pricing speed by directly regulating the time between updates or by limiting the frequency that comparison prices are updated.[26] A platform that offers automated pricing can, by curating the menu of available rules or inserting delays in activating a new rule, give such rules a degree of persistence.[27] Pricing capabilities can differ across sellers on the same platform, either because sellers make different adoption decisions or because the platform gives preferred access to certain sellers.[28] In particular, vertically integrated platforms may use one pricing system for their own products while providing different systems to third-party sellers.[29]

In this section, we present a simple model in which a platform chooses the pricing technology used by two competing sellers, where seller a is the platform's retail arm or a preferred seller. We treat speed and commitment not as exogenous primitives, but as reduced-form outcomes reflecting a platform's preferences. The simple framework allows us to analyze how a platform's design decisions—and the emergence of coercive outcomes—vary with the weights placed on consumer surplus and seller profits.

[26]For example, Wieting and Sapi (2021) report that most repricers on Amazon and Dutch online retailer Bol appear to update prices roughly every 20 minutes. European e-commerce platform Allegro offers an automatic-pricing rule that uses a comparison price that is checked once an hour and updates the seller's offer within that cycle (Allegro, 2026).

[27]For example, Amazon's seller documentation states that Automate Pricing typically processes price updates under existing rules in less than 15 minutes, while new rules or changes in rule parameters can take up to one hour to apply to all assigned products (Amazon, 2026). Other marketplaces impose analogous constraints. For example, German platform Kaufland states that automated price-adjustment rules can take up to 30 minutes to activate, and that switching from "Smart Pricing" to "Smart Pricing Plus" can take about two days to first activate (Kaufland Global Marketplace, 2026a,b).

[28]The Dutch e-commerce platform Bol states that higher API rate limits are available only in limited cases for exceptionally large partners (bol.com, 2026). eBay states that default API limits are designed for individuals and smaller businesses and that higher limits require a vetting process (eBay, 2026). Walmart states that API "[r]ate limits are allotted at the seller level, with each seller receiving their own specific limits." (Walmart, 2026).

[29]The European Commission's Amazon Marketplace investigation provides an example. The Commission expressed the concern that Amazon Retail used non-public data from third-party sellers' marketplace activity to inform retail decisions, including the prices of its own offers (European Commission, 2022b).

6.1 Platform Objective

Assume two sellers (or firms), 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 is the relevant limit when periods are short.[30] Seller b does not use an algorithm and instead chooses prices to maximize per-period profits as in Assumption A1, while seller a is restricted to best-response punishments as in Assumption A2.

We focus on the symmetric linear demand system in equation (19). 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 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 payoffs increase with each seller's profit. This formulation is most natural for settings in which the platform's payoff is directly aligned with seller profits and could arise from vertical integration or preferred-seller arrangements. Under our zero-marginal-cost normalization, profits and seller revenues (gross merchandise value) are the same, so κ_a and κ_b could equivalently be interpreted as percentage commissions on revenues.[31] 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

$$\max_{\alpha \in [0,1], \gamma \in [0,1]} \; \kappa_{CS}\, CS^*(\alpha, \gamma) \; + \; \kappa_a\, \pi_a^*(\alpha, \gamma) \; + \; \kappa_b\, \pi_b^*(\alpha, \gamma). \tag{32}$$

Two comparative statics from earlier results discipline this problem. First, under linear demand (Assumption D1), consumer surplus is decreasing in both speed and commitment (Proposition 7). Second, the algorithmic seller's profit is increasing in both speed and commitment (Lemma 3). Together, these imply a natural tension: higher (α, γ) benefits seller a and harms consumers, while the effect on seller b can be ambiguous. Thus, a platform's optimal policy depends on the weights $(\kappa_a, \kappa_b, \kappa_{CS})$.

6.2 Platform Incentives

Given our demand system, we can characterize the optimal technological choices of the platform. When maximizing total welfare, the platform chooses simultaneous pricing: no speed

[30] 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.

[31] With zero marginal costs, $\pi_i(p_i, p_{-i}) = p_i D_i(p_i, p_{-i})$. Thus, if the platform charges an ad valorem fee τ_i, its revenue from seller i is $\tau_i p_i D_i(p_i, p_{-i})$, so a revenue objective is equivalent to the profit-weighted objective in (32), after rescaling the weights. With positive marginal costs, the objectives differ because $p_i D_i(p_i, p_{-i}) = \pi_i(p_i, p_{-i}) + c_i D_i(p_i, p_{-i})$, so a transaction-value-based platform objective places additional weight on volume.

advantage and no commitment. If the weight on consumer surplus is sufficiently low, the platform will choose positive values for α and/or γ. If the platform puts increasingly high weight on firm a's profits, it pushes the technology toward maximal speed and full commitment. We collect these results in the following proposition:

Proposition 11 (Platform-Optimal Technology)**.** *In the platform problem described above, with $\beta \to 1$ and demand given by Assumption D1, the following hold:*

(i) If the platform weights consumer surplus and each seller's profit equally, $\kappa_{CS} = \kappa_a = \kappa_b > 0$, then the platform optimally chooses $(\alpha^*, \gamma^*) = (0, 0)$.

(ii) For any (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 $(\alpha^*, \gamma^*) \neq (0, 0)$.

(iii) Fix $\kappa_{CS} \geq 0$ and $\kappa_b \geq 0$. As the platform places increasingly large weight on firm a's profit relative to consumer surplus and firm b's profit, its optimal choice approaches the maximal-speed, full-commitment policy $(\alpha^*, \gamma^*) = (1, 1)$.

Within the class of outcomes implementable by changing (α, γ) in our coercion model, a higher relative speed for firm a and stronger commitment move equilibrium prices upward (Proposition 6). For our quasi-linear utility model, choosing $(\alpha, \gamma) = (0, 0)$ collapses the game to the standard simultaneous price-setting benchmark and maximizes welfare. Thus, a platform that aims to maximize total welfare will not give one seller faster access to price information or faster repricing capability than its rivals.

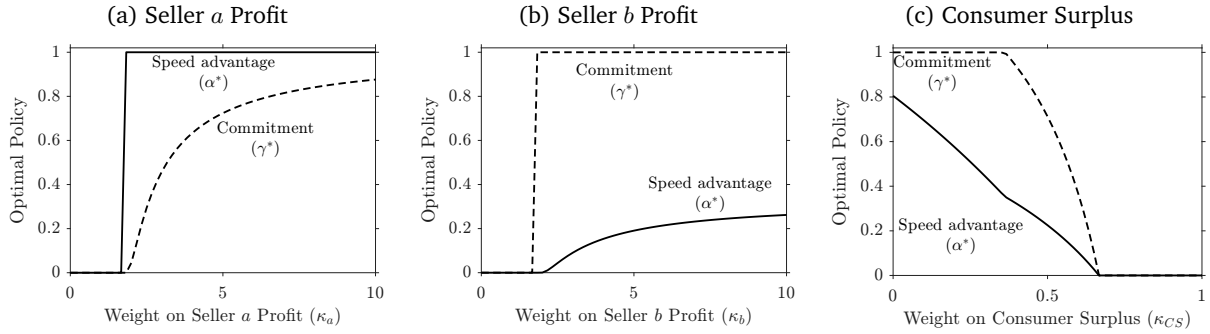
Part (ii) follows by comparing maximal coercion with simultaneous pricing. Both sellers earn strictly more at $(\alpha, \gamma) = (1, 1)$ than at $(0, 0)$, so for any (κ_a, κ_b) with $\kappa_a + \kappa_b > 0$, the gain in seller-profit terms is strictly positive. If the platform places sufficiently little weight on consumer surplus, this gain dominates the consumer-surplus loss from moving away from $(0, 0)$. Thus, a platform that prioritizes seller profit will find it optimal to provide one firm with a speed advantage and enable commitment to a pricing rule. To generate supracompetitive outcomes, a platform need not explicitly coordinate prices; it can do so through its technical offerings alone.

Finally, if the platform disproportionately internalizes seller a's profit, part (iii) applies. This can occur if, for example, the platform is vertically integrated with seller a. Under linear demand, $\pi_a^*(\alpha, \gamma)$ is increasing in both arguments (Lemma 3). As the platform's weight on seller a grows, this profit term dominates and the platform chooses a policy arbitrarily close to $(1, 1)$.

Figure 10 illustrates how the platform uses the two instruments. We examine the platform's optimal choice of within-period speed advantage, α^* (solid), and 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 sets $\alpha^* = 1$ and then increases commitment more gradually. Intuitively, speed is key for coercion: without

Figure 10: Optimal Platform Policy as Platform Objectives Vary



Notes: Each panel plots the platform's optimal choice of speed advantage α^* (solid) and commitment parameter γ^* (dashed) solving (32) 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. Values are computed under (19) with $d = 1$.

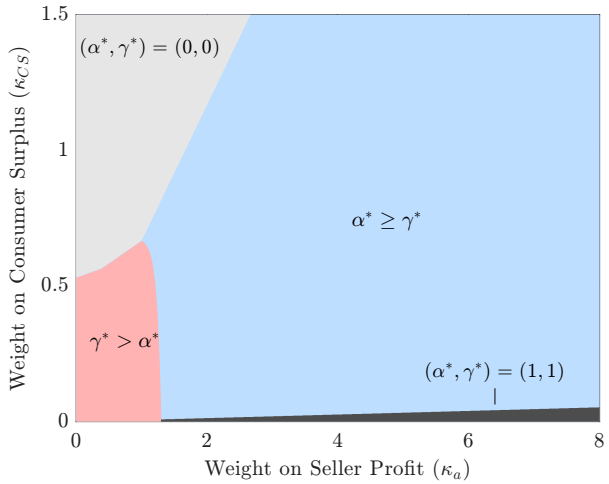
sufficiently fast within-period reactions, punishments arrive too late to strongly reshape seller b's static incentives. Once speed is high, additional commitment further increases seller a's ability to sustain more aggressive target prices. Conversely, as the weight on seller b's profits rises, the platform initially enables full 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 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 speed asymmetries and persistent rules, if possible.

Figure 11 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 10. 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), its optimal choice of technology is close to $(1, 1)$.

Between these extremes, Figure 11 shows two intermediate regimes that help interpret how the platform softens competition away from the two extreme policies. In the blue region, $\alpha^* \geq \gamma^*$ and the platform prioritizes a speed advantage, enabling 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. When the platform can choose these parameters freely, the model suggests that the prioritization of one component over another can be informative about the relative weights in the platform's objective.

Figure 11: 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. The region labeled $(\alpha^*, \gamma^*) = (1, 1)$ denotes the region satisfying $\alpha^* > 0.99$ and $\gamma^* > 0.99$ and therefore represents policies close to $(1, 1)$. Values are computed under equation (19) with $d = 1$.

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, where they can, discourage 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.

These findings complement the literature focused on how platforms and other third parties can facilitate or mitigate coordination through demand steering or information sharing (e.g., Johnson et al., 2023; Harrington, 2022; Sugaya and Wolitzky, 2025). 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 with a fast and persistent pricing algorithm can sustain supracompetitive outcomes. In a coercive equilibrium, prices can be substantially higher than competitive levels and consumer surplus can be lower than under full collusion. This is true even when the algorithmic firm's rivals do not (and cannot) support collusion through forward-looking strategies, sophisticated dynamic reasoning, explicit communication, or common adoption of the same algorithm. Our analysis shows that high prices need not

reflect mutual coordination; a coercive equilibrium can be attained by unilateral action.

The mechanism is straightforward. A speed advantage lets the algorithmic firm punish deviations quickly, reshaping a rival's current-period profit function. Commitment—through the persistence of the deployed pricing rule—then makes a high-price outcome sustainable for the algorithmic firm. Together these features expand the set of price pairs the algorithmic firm can implement relative to standard leader–follower benchmarks.

We showed that this coercive mechanism is robust to several modifications of the environment. It remains an equilibrium when the rival is forward-looking. It weakens but does not disappear in oligopoly settings. Alternative update rules, including price matching and harsher punishments, change the set of implementable targets but preserve the role of speed and commitment. Coercion can arise even when rivals are naive and use simple learning rules based only on their own past prices and profits. Hence, supracompetitive outcomes do not even require rivals to be aware that they face an algorithmic opponent.

Our framework also has implications for platform design. Platforms can influence the pricing speed and commitment of sellers through technical choices about, e.g., latency, data access, and default rules. When a platform places more weight on seller profits than consumer surplus, it has an incentive to equip a subset of sellers with fast, persistent algorithms in order to soften competition. Policies that limit rapid automatic responses to rivals' prices, restrict repricing rules, or require neutral access to pricing technology can therefore affect welfare even without addressing explicit coordination.

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 potentially 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. There may be a role for policy to limit such investments when they are socially wasteful, following analogous arguments made about high-frequency trading (Budish et al., 2015). Remaining challenges for future research include a more thorough treatment of algorithm adoption and investments in speed and commitment, an analysis of richer learning and behavioral rules for slower firms, and 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.

Airbnb (2026). Use Smart Pricing to automatically adjust your prices based on demand. Airbnb Help Center, https://www.airbnb.com/help/article/1168. Accessed June 23, 2026.

Allegro (2026). How the pricing rules based on lowest price on Allegro and top offer work. https://help.allegro.com/en/sell/a/how-the-pricing-rules-based-on-lowest-price-on-allegro-and-top-offer-work-8drjrq7EkT5. Help for sellers. Accessed May 22, 2026.

Amazon (2026). Automate Pricing. https://sell.amazon.com/tools/automate-pricing. Accessed May 21, 2026.

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.

Aquilina, M., E. Budish, and P. O'Neill (2022). Quantifying the high-frequency trading "arms race". *The Quarterly Journal of Economics 137*(1), 493–564.

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*.

Biais, B., T. Foucault, and S. Moinas (2015). Equilibrium fast trading. *Journal of Financial Economics 116*(2), 292–313.

bol.com (2026). API rate limits. https://partnerplatform.bol.com/en/idp/api-rate-limits. Partner Platform documentation. Accessed May 22, 2026.

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.

Competition and Markets Authority (2018, October). Pricing algorithms: Economic working paper on the use of algorithms to facilitate collusion and personalised pricing. Technical Report CMA94, Competition and Markets Authority.

DealHub (2025). Dynamic pricing. `https://dealhub.io/glossary/dynamic-pricing/`. Accessed February 2025.

eBay (2026). API call limits. `https://developer.ebay.com/develop/get-started/api-call-limits`. eBay Developers Program documentation. Accessed May 22, 2026.

European Commission (2022a). Case AT.40703 – Amazon – Buy Box. Commission Decision C(2022) 9442. Summary published in OJ C 87, 10.3.2023, pp. 11–15.

European Commission (2022b). Cases AT.40703 – Amazon – Buy Box and AT.40462 – Amazon Marketplace. Commission Notice 2022/C 278/06.

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.

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.

Kaufland Global Marketplace (2026a). Automated price adjustments. `https://www.kauflandglobalmarketplace.com/en/seller-university/selling/price-adjustments/`. Seller University documentation. Accessed May 22, 2026.

Kaufland Global Marketplace (2026b). Smart Pricing. `https://www.kauflandglobalmarketplace.com/en/seller-university/selling/smart-pricing/`. Seller University documentation. Accessed May 22, 2026.

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.

Lighthouse (2026). Lighthouse Pricing. https://www.mylighthouse.com/platform/pricing. Accessed June 23, 2026.

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.

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.

Stamatopoulos, I., A. Bassamboo, and A. Moreno (2021). The effects of menu costs on retail performance: Evidence from adoption of the electronic shelf label technology. *Management Science 67*(1), 242–256.

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

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

Walmart (2026). Rate limiting. https://developer.walmart.com/us-marketplace/docs/rate-limiting. Walmart Marketplace API Documentation. Accessed May 21, 2026.

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 best deviation payoff and the set of incentive-compatible prices. 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 and consider a candidate target pair (p_a, p_b). Firm b receives $\pi_b(p_b, p_a)$ if it complies and chooses p_b. Under Assumption A2, if firm b chooses a price $\hat{p}_b$ different from the target 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. Its best deviation payoff, allowing for deviations arbitrarily close to the target, is

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

Because P_b is an interval and the deviation payoff is continuous, this maximum equals the supremum over off-target prices $P_b \setminus \{p_b\}$. Thus, compliance is optimal if and only if $\pi_b(p_b, p_a) \geq D(p_a; \alpha)$. Define the incentive-compatible set

$$\mathcal{F}(p_a; \alpha) \equiv \{p_b \in P_b : \ \pi_b(p_b, p_a) \geq D(p_a; \alpha)\}. \tag{A.2}$$

Denote the largest incentive-compatible 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

$$\Psi_a(p_a, p_b; p_{bt}) = (1-\alpha)\pi_a(p_a, p_{bt}) + \alpha\, \pi_a(\sigma(p_{bt}), p_{bt}) + \omega(\delta)\pi_a(p_a, p_b), \tag{A.3}$$

where $\sigma(p_{bt}) = p_a$ if $p_b = p_{bt}$ and $\sigma(p_{bt}) = R_a(p_{bt})$ otherwise. Firm a's initialization price and within-period update coincide when the current rival price equals the target.

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 conditional value function

$$\widehat{\Psi}_a(p_a; p_b, p_{bt}) \equiv (1-\alpha)\pi_a(p_a, p_{bt}) + \alpha\, \pi_a(\sigma(p_{bt}), p_{bt}) + \omega(\delta)\pi_a(p_a, \bar{p}_b(p_a; \alpha)), \tag{A.4}$$

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 $\sup_{p_{bt} \in P_b} \left| \frac{dp_b^\dagger}{dp_{bt}} \right| < 1$ if $\delta > 0$. If $\delta = 0$, this mapping has a unique fixed point.

(iii) For fixed p_a and p_{bt}, when firm a compares incentive-compatible targets $p_b \in \mathcal{F}(p_a; \alpha)$, it does not consider changes in p_b to alter the current-period compliance of p_{bt}. When $\delta = 0$, ties are resolved in favor of the largest incentive-compatible target.

These conditions admit natural economic interpretations. Condition (i) requires firm a's conditional value function to be strictly quasi-concave along the largest incentive-compatible boundary $\bar{p}_b(\cdot; \alpha)$. This condition implies a unique optimizer $p_a^\dagger(p_{bt})$ when the target for firm b is chosen to satisfy $p_b^\dagger(p_{bt}) = \bar{p}_b(p_a^\dagger(p_{bt}); \alpha)$. Condition (ii) is a contraction requirement that rules out excessive strategic feedback and ensures a unique fixed point. (The case when $\delta = 0$ ensures there is a well-defined selection rule for $p_b^\dagger$, which has no effect on future payoffs under zero commitment.) Demand systems with strictly concave own-price profits and cross-price effects sufficiently small relative to own-price effects can satisfy these conditions. In the linear-demand example used throughout the paper, these requirements can be verified directly. More generally, they can be checked analytically or numerically for a given demand system.

Under condition (iii), firm a's objective is increasing in p_b over $\mathcal{F}(p_a; \alpha)$ (by substitutes), so it optimally selects the largest incentive-compatible target $p_b = \bar{p}_b(p_a; \alpha)$. Specifying how firm a considers compliance is necessary because, absent such a convention, firm a may consider a small deviation downward in p_b when $p_{bt} = \bar{p}_b(p_a; \alpha)$ to be profitable. Because our baseline $\sigma(\cdot)$ treats the event $p_b = p_{bt}$ differently from $p_b \neq p_{bt}$, reducing p_b to reclassify p_{bt} as non-compliant (even when firm b desires to comply) can discontinuously improve the α-weighted term while reducing the continuation payoff by an arbitrarily small amount. Condition (iii) rules out this incentive.

We impose this assumption to ensure intuitive economic behavior, and we view it as natural because firm a selects its algorithm without observing firm b's contemporaneous choice of p_{bt}. In our baseline model, it is unnecessary in the high-commitment limit $\delta = \beta\gamma \to 1$, where the current-period payoff receives vanishing weight relative to the continuation payoff. It is also unnecessary with marginal-cost punishment, where firm a's punishment payoff is zero. This discontinuity issue can alternatively be avoided by augmenting the model with a discrete price grid or a small compliance band around the target price, by imposing a vanishing cost of triggering within-period punishment, or by invoking alternative timing assumptions, as we discuss in Appendix A.2.

Proof. Fix $p_{bt} \in P_b$ and let (p_a, p_b) denote any optimal target pair satisfying the maintained assumptions. For any choice of p_a, Assumption R1(iii) and the fact that products are substitutes imply that firm a selects the largest incentive-compatible target for firm b, $p_b = \bar{p}_b(p_a; \alpha)$. Thus,

for any optimal pair, p_a must maximize $\widehat{\Psi}_a(p_a; p_b, p_{bt})$. By Assumption R1(i), this function is continuous and strictly quasi-concave on a compact interval. Hence the maximizer $p_a^\dagger(p_{bt})$ exists and 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

$$p_{bt} = p_b^\dagger(p_{bt}). \tag{A.5}$$

If $\delta > 0$, Assumption R1(ii) implies that the mapping $p_{bt} \mapsto p_b^\dagger(p_{bt})$ is continuously differentiable with derivative uniformly 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 $p_{bt}^* = p_b^\dagger(p_{bt}^*)$. If $\delta = 0$, existence and uniqueness of the fixed point follow directly from the second part of Assumption R1(ii). Thus the equilibrium price p_{bt}^* exists and is unique.

Following the result from the first part of the proof, there exists a unique selected target pair $(p_a^*, p_b^*) = \big(p_a^\dagger(p_{bt}^*), p_b^\dagger(p_{bt}^*)\big)$. The fixed-point condition implies $p_b^* = p_{bt}^*$. These prices satisfy the coercive-equilibrium conditions in Definition 1. Because both the fixed point and the selected target pair conditional on that fixed point are unique, no second target pair satisfying the maintained assumptions can satisfy the equilibrium conditions. Therefore, the coercive equilibrium exists and is unique. □

Proof of Lemma 1

Proof. Write the commitment parameter as $\delta \equiv \beta\gamma \in [0, 1)$. In a slight abuse of notation, we write $\delta = 1$ for the full-commitment limit, formally $\delta \to 1$. We prove the four cases in turn. Uniqueness of equilibrium in each case follows from Proposition 1 under the maintained regularity conditions.

For convenience, we reproduce the value function from (10):

$$\tilde{V}(t) = \max_{(p_a^\dagger, p_b^\dagger)} \Big[(1-\alpha)\,\pi_a(\rho, p_{bt}) + \alpha\,\pi_a(\sigma(p_{bt}), p_{bt}) + \frac{\delta}{1-\delta}\,\pi_a(p_a^\dagger, p_b^\dagger)\Big], \tag{A.6}$$

with equilibrium constraints $\rho = p_a^\dagger$ and $\sigma(p_b) = \rho$ if $p_b = p_b^\dagger$ and $\sigma(p_b) = R_a(p_b)$ if $p_b \neq p_b^\dagger$.

No Speed and No Commitment (Simultaneous Pricing). Set $\alpha = 0$ and $\delta = 0$. The algorithm has no within-period update and no continuation value. Firm a therefore solves $\max_\rho \pi_a(\rho, p_b)$ while firm b solves $\max_{p_b} \pi_b(p_b, \rho)$. Thus equilibrium prices satisfy $p_a = R_a(p_b)$ and $p_b = R_b(p_a)$, which is exactly the Bertrand–Nash price vector (p_a^B, p_b^B).

Commitment Only. Set $\alpha = 0$ and allow $\delta \in (0, 1)$. Since the algorithm has no within-period update, firm a charges $p_a = \rho$ throughout each period. Firm b's static problem reduces to

$\max_{p_b} \pi_b(p_b, \rho)$, so any continuation price chosen by firm b must satisfy $p_b = R_b(\rho)$. Hence the equilibrium price vector lies on firm b's best-response function.

At a reprogramming date, firm a chooses ρ taking the current rival price p_{bt} as given, but it internalizes the fact that, if the algorithm remains in force, firm b will play $R_b(\rho)$ in future periods. Thus, firm a solves

$$\max_{\rho \in P_a} \left\{ \pi_a(\rho, p_{bt}) + \frac{\delta}{1-\delta} \pi_a(\rho, R_b(\rho)) \right\}. \tag{A.7}$$

After multiplying by $1-\delta$ and taking $\delta \to 1$, the limiting problem is $\max_\rho \pi_a(\rho, R_b(\rho))$. This is equivalent to the Stackelberg pricing problem in which firm a is the leader. Hence, the equilibrium with full commitment and no speed advantage converges to (p_a^L, p_b^F).

Speed Only. Set $\delta = 0$ and allow $\alpha \in (0, 1)$. Firm a's problem is purely within-period:

$$\max_{\mathcal{A}} (1-\alpha)\pi_a(\rho, p_{bt}) + \alpha \pi_a(\sigma(p_{bt}), p_{bt}). \tag{A.8}$$

For any realized p_{bt}, it is weakly dominant for firm a to set the update rule equal to its static best response, $\sigma(p_{bt}) = R_a(p_{bt})$, and this behavior is maintained by Assumption A2. Setting $\alpha = 1$ yields $\max_{\mathcal{A}} \pi_a(R_a(p_{bt}), p_{bt})$. Assumption A2 requires $\sigma(p_b^\dagger) = \rho$. Choose $\rho = R_a(p_b^\dagger)$, which maximizes firm a's current payoff at the target; for every off-target price, the same assumption gives $\sigma(p_b) = R_a(p_b)$.

Firm b then solves $\max_{p_b} \pi_b(p_b, R_a(p_b))$. This is equivalent to the Stackelberg problem when firm a is the follower. Therefore, the maximal-speed, no-commitment benchmark is the sequential outcome in which firm b is the leader and firm a is the follower.

Maximal Coercion. Finally, set $\alpha = 1$ and take the full-commitment limit $\delta \to 1$. Multiplying firm a's objective by $1-\delta$ leaves the maximizer unchanged, and the normalized objective converges to

$$\max_{(p_a^\dagger, p_b^\dagger)} \pi_a(p_a^\dagger, p_b^\dagger), \tag{A.9}$$

subject to firm b's incentive-compatibility constraint. The absence of best-response replies from the objective means that neither equilibrium price need lie on a static best-response function. Intuitively, the maximal coercion case breaks the usual leader-follower pricing logic; instead, firm a offers two prices to firm b that are incentive compatible.

If firm b complies with the target pair $(p_a^\dagger, p_b^\dagger)$, its payoff is $\pi_b(p_b^\dagger, p_a^\dagger)$, while any deviation receives $\pi_b\big(\hat{p}_b, R_a(\hat{p}_b)\big)$, reflecting the instantaneous reaction by firm a. Therefore, compliance is optimal if and only if

$$\pi_b(p_b^\dagger, p_a^\dagger) \geq \max_{\hat{p}_b} \pi_b\big(\hat{p}_b, R_a(\hat{p}_b)\big). \tag{A.10}$$

By the maintained tie-breaking convention, firm b complies when this constraint holds with equality. Combining firm a's objective with this constraint yields equation (14).

The Stackelberg prices where a is a follower satisfy the constraint with equality. Thus these prices are implementable. But the feasible set of implementable prices can include substantially higher prices, as we characterize in Sections 3 and 4 in the main text.

□

Proof of Proposition 2

Proof. Fix a joint-profit-maximizing vector (p_a^C, p_b^C) satisfying (18). Consider the baseline algorithm with this collusive target, $(p_a^\dagger, p_b^\dagger) = (p_a^C, p_b^C)$ such that $\rho = p_a^C$, $\sigma(p_b^C) = p_a^C$, and $\sigma(\hat{p}_b) = R_a(\hat{p}_b)$ for $\hat{p}_b \neq p_b^C$ (Assumption A2).

If firm b complies, its current-period payoff is $\pi_b(p_b^C, p_a^C)$. If it deviates to $\hat{p}_b \neq p_b^C$, its payoff is $(1-\alpha)\pi_b(\hat{p}_b, p_a^C) + \alpha\pi_b(\hat{p}_b, R_a(\hat{p}_b))$. By Assumption A1 and the tie-breaking convention, the collusive target is implementable at speed α if and only if

$$(1-\alpha)\pi_b(\hat{p}_b, p_a^C) + \alpha\pi_b\big(\hat{p}_b, R_a(\hat{p}_b)\big) \leq \pi_b(p_b^C, p_a^C) \qquad \text{for every } \hat{p}_b \in P_b \setminus \{p_b^C\}. \tag{A.11}$$

First consider deviations where $\pi_b(\hat{p}_b, p_a^C) \leq \pi_b(p_b^C, p_a^C)$, that is, deviations that are immediately unprofitable. By condition (18), $\pi_b(\hat{p}_b, R_a(\hat{p}_b)) \leq \pi_b(p_b^C, p_a^C)$ as well. The left-hand side of (A.11) is a convex combination of these two payoffs, so it is bounded above by $\pi_b(p_b^C, p_a^C)$ for every $\alpha \in [0,1]$.

Next consider deviations where $\pi_b(\hat{p}_b, p_a^C) > \pi_b(p_b^C, p_a^C)$, that is, deviations that would be profitable if firm a did not respond within the period. For any such $\hat{p}_b$, condition (18) gives

$$\pi_b\big(\hat{p}_b, R_a(\hat{p}_b)\big) \leq \pi_b(p_b^C, p_a^C) < \pi_b(\hat{p}_b, p_a^C). \tag{A.12}$$

Given this relation, the left-hand side of (A.11) is strictly decreasing in α. We can rewrite inequality (A.11) as

$$\alpha \;\geq\; \frac{\pi_b(\hat{p}_b, p_a^C) - \pi_b(p_b^C, p_a^C)}{\pi_b(\hat{p}_b, p_a^C) - \pi_b\big(\hat{p}_b, R_a(\hat{p}_b)\big)} \equiv \alpha(\hat{p}_b). \tag{A.13}$$

Since the numerator is positive and weakly smaller than the denominator, $\alpha(\hat{p}_b) \in (0,1]$ for every such deviation. Define

$$\alpha^C \equiv \sup\big\{\alpha(\hat{p}_b) : \hat{p}_b \in P_b \setminus \{p_b^C\},\ \pi_b(\hat{p}_b, p_a^C) > \pi_b(p_b^C, p_a^C)\big\}, \qquad \sup\emptyset \equiv 0, \tag{A.14}$$

so that $\alpha^C \in [0,1]$. Because the left-hand side of (A.11) is strictly decreasing in α for such deviations, (A.11) holds for every deviation for any $\alpha \in [\alpha^C, 1]$.

Thus, for $\alpha \in [\alpha^C, 1]$, firm b complies and chooses p_b^C and firm a charges p_a^C throughout the period, implementing the collusive price vector. Conversely, for any $\alpha < \alpha^C$ there is a deviation

$\hat{p}_b$ with $\alpha(\hat{p}_b) > \alpha$, at which (A.11) fails. Thus, the set of α that can implement the collusive price vector is exactly $[\alpha^C, 1]$. □

Proof of Proposition 3

Proof. Fix an arbitrary commitment level $\delta \equiv \beta\gamma \in [0, 1)$, and let

$$\mathcal{E} \equiv \left\{ \left(p_a^\dagger(\tilde{\alpha}), p_b^\dagger(\tilde{\alpha})\right) : \tilde{\alpha} \in [0, 1] \right\} \tag{A.15}$$

denote the set of equilibrium target pairs across all speed levels, holding δ fixed. At speed α, call a pair $(p_a, p_b) \in \mathcal{E}$ securable if it satisfies firm b's incentive constraint,

$$\pi_b(p_b, p_a) \geq D(p_a; \alpha). \tag{A.16}$$

Firm a's securable per-period profit at speed α is the supremum of $\pi_a(p_a, p_b)$ over these pairs. The set over which the supremum is taken is nonempty because the equilibrium target pair at speed α belongs to $\mathcal{E}$ and is incentive compatible.

Consider two speed levels $\alpha' > \alpha$, and let $(p_a, p_b) \in \mathcal{E}$ be any target pair that is securable at speed α. Let $\hat{p}_b$ attain firm b's best deviation payoff $D(p_a; \alpha')$ at speed α'. Because $\hat{p}_b$ is an optimal deviation from an equilibrium target pair, the condition in the proposition implies

$$\pi_b\big(\hat{p}_b, R_a(\hat{p}_b)\big) \leq \pi_b(\hat{p}_b, p_a). \tag{A.17}$$

Using this inequality and the definition of D,

$$\begin{aligned} D(p_a; \alpha') &= (1-\alpha)\pi_b(\hat{p}_b, p_a) + \alpha\pi_b\big(\hat{p}_b, R_a(\hat{p}_b)\big) \\ &\quad + (\alpha' - \alpha)\left[\pi_b\big(\hat{p}_b, R_a(\hat{p}_b)\big) - \pi_b(\hat{p}_b, p_a)\right] \\ &\leq (1-\alpha)\pi_b(\hat{p}_b, p_a) + \alpha\pi_b\big(\hat{p}_b, R_a(\hat{p}_b)\big) \\ &\leq D(p_a; \alpha). \end{aligned} \tag{A.18}$$

The last inequality follows because $D(p_a; \alpha)$ is firm b's best deviation payoff at speed α.

Since (p_a, p_b) is securable at speed α,

$$\pi_b(p_b, p_a) \geq D(p_a; \alpha) \geq D(p_a; \alpha'). \tag{A.19}$$

Thus the same target pair is securable at speed α'. Because (p_a, p_b) was arbitrary, every pair in $\mathcal{E}$ that is securable at speed α remains securable at speed α'. Therefore,

$$\sup_{\substack{(p_a,p_b)\in\mathcal{E} \\ \pi_b(p_b,p_a)\geq D(p_a;\alpha')}} \pi_a(p_a, p_b) \geq \sup_{\substack{(p_a,p_b)\in\mathcal{E} \\ \pi_b(p_b,p_a)\geq D(p_a;\alpha)}} \pi_a(p_a, p_b). \tag{A.20}$$

Hence firm a's securable per-period profit is weakly increasing in the speed advantage α.

The result concerns the highest profit that firm a can secure from among the equilibrium target pairs in $\mathcal{E}$. It does not assert that equilibrium profit is weakly increasing without further restrictions, and it does not require the punishment condition to hold for target pairs outside $\mathcal{E}$. □

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 target pair $(p_a^\dagger, p_b^\dagger)$ to maximize (10) subject to the implementation constraints (i)–(ii) and firm b's incentive constraint (iii). Define

$$G(p_a^\dagger, p_b^\dagger, \rho, \sigma; \delta) \equiv (1-\alpha)\pi_a(\rho, p_{bt}) + \alpha\pi_a(\sigma(p_{bt}), p_{bt}) + \frac{\delta}{1-\delta}\,\pi_a(p_a^\dagger, p_b^\dagger), \tag{A.21}$$

which is the value function conditional on δ. We suppress p_{bt} in the notation for G.

Take $\delta' > \delta$. Let $(p_a^\dagger, p_b^\dagger, \rho, \sigma)$ solve the problem at δ, and let $(\tilde{p}_a^\dagger, \tilde{p}_b^\dagger, \tilde{\rho}, \tilde{\sigma})$ solve the problem at δ'. Note that the feasible set of incentive-compatible prices does not depend on δ. Since the feasible set is the same at δ and δ', optimality implies $G(p_a^\dagger, p_b^\dagger, \rho, \sigma; \delta) \geq G(\tilde{p}_a^\dagger, \tilde{p}_b^\dagger, \tilde{\rho}, \tilde{\sigma}; \delta)$ and $G(\tilde{p}_a^\dagger, \tilde{p}_b^\dagger, \tilde{\rho}, \tilde{\sigma}; \delta') \geq G(p_a^\dagger, p_b^\dagger, \rho, \sigma; \delta')$. Adding these two inequalities and canceling out the current-period payoff terms yields

$$\left(\frac{\delta'}{1-\delta'} - \frac{\delta}{1-\delta}\right)\left[\pi_a(\tilde{p}_a^\dagger, \tilde{p}_b^\dagger) - \pi_a(p_a^\dagger, p_b^\dagger)\right] \geq 0. \tag{A.22}$$

Because $\delta' > \delta$, the first term is positive. Hence $\pi_a(\tilde{p}_a^\dagger, \tilde{p}_b^\dagger) \geq \pi_a(p_a^\dagger, p_b^\dagger)$. Thus firm a's per-period profit at the target price pair, conditional on the current rival price, is weakly increasing in effective commitment $\beta\gamma$.

□

Proof of Proposition 5

Proof. Write $\delta \equiv \beta\gamma$ and recall the best deviation payoff $D(p_a; \alpha)$, the incentive-compatible set $\mathcal{F}(p_a; \alpha)$, the maximal incentive-compatible target $\bar{p}_b(p_a; \alpha)$, the continuation weight $\omega(\delta) = \delta/(1-\delta)$, and firm a's objective $\Psi_a(p_a, p_b; p_{bt})$, all defined before the proof of Proposition 1.

First, we show that the collusive target satisfies firm b's incentive constraint with slack in the maximal-speed limit. At $\alpha = 1$, the best deviation payoff does not depend on firm a's target price and equals $D(p_a; 1) = \max_{\hat{p}_b \in P_b} \pi_b\big(\hat{p}_b, R_a(\hat{p}_b)\big) = \pi_b(p_b^L, p_a^F)$. By the maintained condition in the proposition, $\pi_b(p_b^C, p_a^C) > \pi_b(p_b^L, p_a^F)$.

Since the collusive price vector is interior and satisfies firm b's incentive constraint with

strict slack, we can therefore choose $\varepsilon > 0$ small enough that $p_b^C + \varepsilon \in P_b$ and $\pi_b(p_b^C + \varepsilon, p_a^C) > \pi_b(p_b^L, p_a^F)$. Since $D(p_a^C; \alpha)$ is continuous in α and equals $\pi_b(p_b^L, p_a^F)$ at $\alpha = 1$, there exists $\overline{\alpha} < 1$ such that, for every $\alpha > \overline{\alpha}$, $\pi_b(p_b^C + \varepsilon, p_a^C) \geq D(p_a^C; \alpha)$ and $p_b^C + \varepsilon \in \mathcal{F}(p_a^C; \alpha)$.

We have established a threshold for α above which a price pair $(p_a^C, p_b^C + \varepsilon)$ is feasible. By the fact that products are substitutes, $\pi_a(p_a^C, p_b^C + \varepsilon) > \pi_a(p_a^C, p_b^C)$. Hence, firm a can choose target prices that deliver profits strictly above its full collusion profits.

Finally, we show that firm a chooses a target with flow profit above the collusive level when commitment is sufficiently high. In the objective Ψ_a, the current-period term is bounded uniformly: because prices lie in compact sets and π_a is continuous, there exists a value $M < \infty$ such that this term differs by at most M across feasible targets and current prices $p_{bt} \in P_b$.

Consider the continuation term. Target prices $(p_a^C, p_b^C + \varepsilon)$ deliver a per-period profit difference $m \equiv \pi_a(p_a^C, p_b^C + \varepsilon) - \pi_a(p_a^C, p_b^C) > 0$. Since firm a selects the largest feasible target, $\bar{p}_b(p_a^C; \alpha)$, and products are substitutes, it must be that $\pi_a(p_a^C, \bar{p}_b(p_a^C; \alpha)) - \pi_a(p_a^C, p_b^C) \geq m$. In the objective function, the continuation value is scaled by $\omega(\delta)$, and $\omega(\delta) \to \infty$ as $\delta \to 1$. Choose $\bar{\delta} < 1$ such that $\omega(\delta)\, m > M$ for all $\delta > \bar{\delta}$.

Now we show that firm a earns at least full collusion profits by contradiction. Fix any $\alpha > \overline{\alpha}$, $\delta > \bar{\delta}$, and $p_{bt} \in P_b$, and consider any feasible target price p_a with $\pi_a\big(p_a, \bar{p}_b(p_a; \alpha)\big) \leq \pi_a(p_a^C, p_b^C)$. Then

$$\Psi_a\big(p_a^C, \bar{p}_b(p_a^C; \alpha); p_{bt}\big) - \Psi_a\big(p_a, \bar{p}_b(p_a; \alpha); p_{bt}\big) \geq \omega(\delta)\, m - M > 0. \tag{A.23}$$

Therefore, no target with $\pi_a\big(p_a, \bar{p}_b(p_a; \alpha)\big) \leq \pi_a(p_a^C, p_b^C)$ is optimal. Because this comparison holds for every current price p_{bt}, it holds in equilibrium. Therefore, the equilibrium target pair satisfies $\pi_a(p_a^*, p_b^*) > \pi_a(p_a^C, p_b^C)$ for $\delta > \bar{\delta}$. □

Proof of Proposition 6

Proof. Write linear demand as

$$D_i(p_i, p_{-i}) = 1 - A p_i + B p_{-i}, \qquad A \equiv \frac{1 + 2d}{4}, \qquad B \equiv \frac{d}{2}. \tag{A.24}$$

Then $A - B = 1/4 > 0$, marginal costs are zero, and

$$\pi_i(p_i, p_{-i}) = p_i(1 - A p_i + B p_{-i}), \qquad R_i(p_{-i}) = \frac{1 + B p_{-i}}{2A}. \tag{A.25}$$

Let $\delta \equiv \beta\gamma$. The main proof is for $(\alpha, \delta) \in (0, 1) \times (0, 1)$. The cases $\alpha = 0$, $\delta = 0$, $\alpha = 1$, and the full-commitment limit can be verified directly from the closed-form boundary solutions and one-sided continuity, and yield the stated comparisons.

We first characterize firm b's optimal deviation, which will pin down the incentive compatibility condition. Fix a candidate target p_a for firm a. If firm b deviates to $\hat{p}_b$, its payoff

is

$$(1-\alpha)\hat{p}_b(1-A\hat{p}_b+Bp_a)+\alpha\hat{p}_b\big(1-A\hat{p}_b+BR_a(\hat{p}_b)\big) \tag{A.26}$$

$$=K(p_a;\alpha)\hat{p}_b-C(\alpha)\hat{p}_b^2, \tag{A.27}$$

where

$$K(p_a;\alpha)\equiv 1+(1-\alpha)Bp_a+\frac{\alpha B}{2A},\qquad C(\alpha)\equiv A-\frac{\alpha B^2}{2A}. \tag{A.28}$$

Because $A>B$ and $\alpha\leq 1$, $C(\alpha)>0$, so the deviation problem is strictly concave with solution

$$\hat{p}_b(p_a;\alpha)=\frac{K(p_a;\alpha)}{2C(\alpha)},\qquad D(p_a;\alpha)=\frac{K(p_a;\alpha)^2}{4C(\alpha)}. \tag{A.29}$$

Firm a selects the largest target price for firm b consistent with firm b's incentive constraint $\pi_b(p_b,p_a)\geq D(p_a;\alpha)$. At an interior solution this constraint binds, so $p_b=\bar{p}_b(p_a;\alpha)$ is the upper root of

$$Ap_b^2-(1+Bp_a)p_b+D(p_a;\alpha)=0. \tag{A.30}$$

Firm a chooses target prices subject to the binding incentive constraint (A.30). Denote the slack as $I_b(p_a,p_b;\alpha)\equiv\pi_b(p_b,p_a)-D(p_a;\alpha)$ and let $\lambda\geq 0$ denote the multiplier on this constraint. Differentiate firm a's constrained problem with respect to its own target and firm b's target, evaluate at the equilibrium fixed point $p_{bt}=p_b$, and multiply both conditions by $1-\delta$ to obtain

$$1-2Ap_a+Bp_b+(1-\delta)\lambda\frac{\partial I_b}{\partial p_a}=0, \tag{A.31}$$

$$\delta Bp_a+(1-\delta)\lambda\frac{\partial I_b}{\partial p_b}=0. \tag{A.32}$$

Here the second equation uses Assumption R1(iii): holding fixed p_a, changes among incentive-compatible targets $p_b\in\mathcal{F}(p_a;\alpha)$ do not alter whether the current price p_{bt} is treated as compliant.

Eliminating λ using $\partial\bar{p}_b/\partial p_a=-(\partial I_b/\partial p_a)/(\partial I_b/\partial p_b)$ yields the implicit equation for firm a's target,

$$H(p_a;\alpha,\delta)\equiv 1-2Ap_a+B\bar{p}_b(p_a;\alpha)+\delta Bp_a\frac{\partial\bar{p}_b}{\partial p_a}(p_a;\alpha)=0. \tag{A.33}$$

It can be verified by direct differentiation of the upper root of (A.30), together with the maintained regularity conditions, that

$$\frac{\partial\bar{p}_b}{\partial p_a}>0,\qquad \frac{\partial\bar{p}_b}{\partial\alpha}>0,\qquad \frac{\partial^2\bar{p}_b}{\partial p_a\,\partial\alpha}>0,\qquad \frac{\partial H}{\partial p_a}<0 \tag{A.34}$$

at the selected interior equilibrium.

Differentiating (A.33) with respect to δ and α yields

$$\frac{\partial H}{\partial \delta} = Bp_a \frac{\partial \bar{p}_b}{\partial p_a} > 0, \qquad \frac{\partial H}{\partial \alpha} = B\frac{\partial \bar{p}_b}{\partial \alpha} + \delta B p_a \frac{\partial^2 \bar{p}_b}{\partial p_a \, \partial \alpha} > 0. \tag{A.35}$$

Implicit differentiation of (A.33) therefore implies, for $\theta \in \{\alpha, \delta\}$,

$$\frac{\partial p_a^\dagger}{\partial \theta} = -\frac{\partial H/\partial \theta}{\partial H/\partial p_a} > 0. \tag{A.36}$$

Since $p_b^\dagger = \bar{p}_b(p_a^\dagger; \alpha)$,

$$\frac{\partial p_b^\dagger}{\partial \delta} = \frac{\partial \bar{p}_b}{\partial p_a}\frac{\partial p_a^\dagger}{\partial \delta} > 0, \qquad \frac{\partial p_b^\dagger}{\partial \alpha} = \frac{\partial \bar{p}_b}{\partial \alpha} + \frac{\partial \bar{p}_b}{\partial p_a}\frac{\partial p_a^\dagger}{\partial \alpha} > 0. \tag{A.37}$$

Since $\delta = \beta\gamma$, both equilibrium target prices are strictly increasing in speed and effective commitment. □

Following this result, we can establish the monotonicity of equilibrium per-period profits.

Lemma 3 (Equilibrium Profit Monotonicity)**.** *Under Assumption D1, the algorithmic firm's equilibrium per-period profit is strictly increasing in speed and effective commitment $\delta \equiv \beta\gamma$ on $[0,1] \times [0,1)$ and is uniquely maximized at the limit $(\alpha, \delta) = (1,1)$.*

Proof. Write $\pi_a^*(\alpha, \delta) = \pi_a\big(p_a^\dagger, \bar{p}_b(p_a^\dagger; \alpha)\big)$. For $\theta \in \{\alpha, \delta\}$,

$$\frac{d\pi_a^*}{d\theta} = \left[1 - 2Ap_a^\dagger + Bp_b^\dagger + Bp_a^\dagger \frac{\partial \bar{p}_b}{\partial p_a}\right]\frac{\partial p_a^\dagger}{\partial \theta} + Bp_a^\dagger \frac{\partial \bar{p}_b}{\partial \theta}, \tag{A.38}$$

where $\partial \bar{p}_b/\partial \delta = 0$. At a solution of (A.33),

$$1 - 2Ap_a^\dagger + Bp_b^\dagger = -\delta B p_a^\dagger \frac{\partial \bar{p}_b}{\partial p_a}, \tag{A.39}$$

and therefore, since $\delta < 1$ and $p_a^\dagger > 0$,

$$1 - 2Ap_a^\dagger + Bp_b^\dagger + Bp_a^\dagger \frac{\partial \bar{p}_b}{\partial p_a} = (1-\delta) B p_a^\dagger \frac{\partial \bar{p}_b}{\partial p_a} > 0. \tag{A.40}$$

It follows that

$$\frac{\partial \pi_a^*}{\partial \delta} = (1-\delta) B p_a^\dagger \frac{\partial \bar{p}_b}{\partial p_a}\frac{\partial p_a^\dagger}{\partial \delta} > 0 \tag{A.41}$$

and

$$\frac{\partial \pi_a^*}{\partial \alpha} = (1-\delta) B p_a^\dagger \frac{\partial \bar{p}_b}{\partial p_a}\frac{\partial p_a^\dagger}{\partial \alpha} + Bp_a^\dagger \frac{\partial \bar{p}_b}{\partial \alpha} > 0. \tag{A.42}$$

Thus firm a's equilibrium per-period profit is strictly increasing in speed and effective commitment on $(\alpha, \delta) \in (0,1) \times (0,1)$. The cases $\alpha = 0$, $\delta = 0$, $\alpha = 1$, and the full-commitment limit can be verified directly from the closed-form frontier and boundary solutions. □

Proof of Proposition 7

Proof. Let CS denote consumer surplus. In the economically relevant region, $D_i(p_i, p_{-i}) \geq 0$ for each firm. By Roy's identity,

$$\frac{\partial CS(p_a, p_b)}{\partial p_i} = -D_i(p_i, p_{-i}) \leq 0 \qquad \text{for } i \in \{a, b\}. \tag{A.43}$$

Thus consumer surplus is weakly decreasing in each price.

Moreover, at equilibrium target prices, firm b's incentive constraint binds at a positive payoff, implying $D_b(p_b^\dagger, p_a^\dagger) > 0$, and firm a earns at least its Bertrand payoff (Lemma 3), implying $D_a(p_a^\dagger, p_b^\dagger) > 0$. Thus, the derivatives of consumer surplus with respect to prices are strictly negative at equilibrium target prices. By Proposition 6, both equilibrium target prices $p_a^\dagger$ and $p_b^\dagger$ are strictly increasing in α and in $\beta\gamma$. Therefore, equilibrium consumer surplus $CS(p_a^\dagger, p_b^\dagger)$ is strictly decreasing in α and in $\beta\gamma$. □

Proof of Proposition 8

Proof. Fix any target pair $(p_a^\dagger, p_b^\dagger)$ and define firm b's best deviation payoff under $\sigma_\eta(\cdot)$ as the supremum over off-target deviations,

$$D(p_a^\dagger, p_b^\dagger; \eta) \equiv \sup_{\hat{p}_b \in P_b \setminus \{p_b^\dagger\}} \Big\{ (1-\alpha)\pi_b(\hat{p}_b, p_a^\dagger) + \alpha \pi_b\big(\hat{p}_b, \sigma_\eta(\hat{p}_b)\big) \Big\}. \tag{A.44}$$

Now let $\eta' \geq \eta$. For every $\hat{p}_b \in P_b \setminus \{p_b^\dagger\}$,

$$\sigma_\eta(\hat{p}_b) - \sigma_{\eta'}(\hat{p}_b) = (\eta' - \eta)\big(R_a(\hat{p}_b) - \underline{p}_a\big) \geq 0, \tag{A.45}$$

because $R_a(\hat{p}_b) \in P_a$ and $\underline{p}_a = \min P_a$. Because firm b's profit is weakly increasing in firm a's price (by substitutes), it follows that

$$\pi_b\big(\hat{p}_b, \sigma_{\eta'}(\hat{p}_b)\big) \leq \pi_b\big(\hat{p}_b, \sigma_\eta(\hat{p}_b)\big) \qquad \text{for all } \hat{p}_b \neq p_b^\dagger. \tag{A.46}$$

Hence $D(p_a^\dagger, p_b^\dagger; \eta') \leq D(p_a^\dagger, p_b^\dagger; \eta)$ whenever $\eta' \geq \eta$. Thus the best deviation payoff is weakly decreasing in η. Since the compliance payoff is independent of η, any target pair implementable under the milder punishment σ_η is also implementable under the harsher punishment $\sigma_{\eta'}$ for $\eta' \geq \eta$. In particular, $\eta = 1$ gives the lowest deviation payoff and therefore the largest implementable set, so the lowest-price punishment provides an upper bound on the coercive power

of the algorithm within this class. □

Proof of Proposition 9

Proof. It is convenient to write demand as

$$D_i(p_i, p_{-i}) = 1 - Ap_i + b\sum_{\ell \neq i} p_\ell, \qquad A \equiv \frac{1}{4} + \frac{d}{2}, \qquad b \equiv \frac{d}{2(N-1)}. \tag{A.47}$$

Fix $K \in \{2, \ldots, N\}$. We first characterize the K-firm cartel benchmark and then ask whether its price vector satisfies every slow rival's incentive constraint.

By symmetry, the K cartel participants ("insiders") charge a common price p_I^K, and the $N - K$ outsiders, when present, charge a common price p_O^K. The cartel price maximizes the profits for insiders, while each outsider chooses its own price noncooperatively. The corresponding first-order conditions are

$$\big(2A - 2b(K-1)\big)p_I^K - b(N-K)p_O^K = 1, \tag{A.48}$$

$$-bKp_I^K + \big(2A - b(N-K-1)\big)p_O^K = 1, \tag{A.49}$$

where (A.49) is omitted when $K = N$. Using the insider first-order condition, demand and profit can be written as

$$D_I^K = \big(A - b(K-1)\big)p_I^K, \qquad \Pi_I^K = \big(A - b(K-1)\big)(p_I^K)^2. \tag{A.50}$$

Now consider the cartel price vector as the coercive target, where firm a and $K - 1$ slow rivals choose p_I^K and $N-K$ outsiders charge p_O^K. Consider first an outsider. Its target price p_O^K is a static best response to the cartel target prices. A deviation triggers marginal-cost punishment, which lowers firm a's price from $p_I^K > 0$ to zero and therefore lowers the outsider's residual demand. In the maximal coercion case, this lowers the payoff to the outsider. Thus every slow outsider's incentive constraint is satisfied.

It remains to check a slow cartel member. If that firm deviates to $\hat{p}_I$, firm a charges marginal cost and every other slow firm remains at its target. The deviating firm earns

$$\hat{p}_I\left[1 - A\hat{p}_I + b\big((K-2)p_I^K + (N-K)p_O^K\big)\right]. \tag{A.51}$$

This is a strictly concave quadratic in $\hat{p}_I$, so its maximum is

$$\frac{\left[1 + b\big((K-2)p_I^K + (N-K)p_O^K\big)\right]^2}{4A}. \tag{A.52}$$

Equation (A.48) implies

$$1 + b\big((K-2)p_I^K + (N-K)p_O^K\big) = (2A - bK)p_I^K. \tag{A.53}$$

Therefore, the slow cartel member complies if and only if

$$\big(A - b(K-1)\big)(p_I^K)^2 \geq \frac{(2A-bK)^2(p_I^K)^2}{4A}. \tag{A.54}$$

Simplifying this inequality gives

$$K^2 \leq \frac{4A}{b} = \frac{2(1+2d)(N-1)}{d}. \tag{A.55}$$

The outsider constraints always hold, so (A.55) is both necessary and sufficient for the K-firm cartel price vector to be implementable. This proves the first statement in part (i). Moreover,

$$\frac{2(1+2d)(N-1)}{d} \geq \frac{2(1+2d)}{d} > 4 \tag{A.56}$$

for every $N \geq 2$ and $d > 0$. Hence the condition holds strictly at $K = 2$, establishing the final statement in part (i).

For part (ii), the largest implementable cartel is the largest integer $K \leq N$ satisfying (A.55):

$$K^*(N,d) = \min\left\{N, \left\lfloor \sqrt{\frac{2(1+2d)(N-1)}{d}} \right\rfloor\right\}. \tag{A.57}$$

The case $N = 2$ is immediate from part (i). For $N \geq 3$, the grand cartel is implementable if and only if

$$N^2 \leq \frac{2(1+2d)(N-1)}{d}, \tag{A.58}$$

which is equivalent to

$$0 < d \leq \frac{2(N-1)}{(N-2)^2}. \tag{A.59}$$

The cutoff is strictly positive for every N. Therefore, sufficiently differentiated products (a sufficiently small d) make the grand cartel implementable. This proves part (ii).

Finally, consider firm a's profit. Whenever a K-firm cartel target is implementable, it is feasible in the maximal-coercion problem and gives firm a the insider profit Π_I^K in (A.50). Because firm a chooses the most profitable implementable target, firm a earns at least Π_I^K and therefore at least the profit associated with $K = K^*(N,d)$.

The comparison is strict whenever (A.55) holds with strict inequality. In that case, the cartel members' incentive constraints are slack. The outsiders' constraints are also slack because marginal-cost punishment strictly lowers their residual-demand intercept. By continuity, firm a

can keep its own target at p_I^K and raise every slow rival's target by a sufficiently small $\varepsilon > 0$ without violating any incentive constraint. Since products are substitutes, this strictly increases firm a's profit. The maximal-coercion profit is therefore strictly greater than Π_I^K whenever the implementability inequality is strict. Applying this argument to $K = 2$ gives the strict comparison with a two-firm cartel. Applying it to $K = N$ gives the strict comparison with the grand cartel whenever $d < 2(N-1)/(N-2)^2$. This proves part (iii). □

Proof of Lemma 2

Proof. Differentiating equation (30) twice with respect to p_b gives

$$\frac{d^2\widetilde{\pi}_b(p_b)}{dp_b^2} = (1-\alpha)\,\frac{\partial^2\pi_b(p_b,\,p_a^\dagger)}{\partial p_b^2} + \alpha\left[\frac{\partial^2\pi_b(p_b,\,\phi p_b)}{\partial p_b^2} + 2\phi\,\frac{\partial^2\pi_b(p_b,\,\phi p_b)}{\partial p_b\,\partial p_a} + \phi^2\,\frac{\partial^2\pi_b(p_b,\,\phi p_b)}{\partial p_a^2}\right]. \tag{A.60}$$

The maintained inequality therefore implies that $\widetilde{\pi}_b$ is strictly concave on the relevant price domain. Any stationary point is consequently its unique maximizer, proving the first two claims.

Under Assumption D1,

$$\pi_b(p_b, p_a) = p_b\left[1 - \left(\tfrac{1}{4} + \tfrac{d}{2}\right)p_b + \tfrac{d}{2}p_a\right]. \tag{A.61}$$

Substituting the linear rule $\sigma(p_b) = \phi p_b$ into equation (30) yields

$$\widetilde{\pi}_b(p_b) = \underbrace{\left[1 + \tfrac{d}{2}(1-\alpha)p_a^\dagger\right]}_{\equiv L} p_b - \underbrace{\left[\tfrac{1}{4} + \tfrac{d}{2} - \tfrac{\alpha d}{2}\phi\right]}_{\equiv \kappa} p_b^2. \tag{A.62}$$

Thus perceived profit is quadratic. For the optimal linear rule, equation (A.117) gives

$$\phi = \frac{(2d+1)(3d+1)}{(\alpha+5)d(d+1)+1}. \tag{A.63}$$

Substitution into the quadratic coefficient gives

$$\kappa = \frac{(2d+1)\big[(5d^2+5d+1) - \alpha(5d^2+d)\big]}{4\big[(\alpha+5)d(d+1)+1\big]} \geq \frac{(2d+1)(4d+1)}{4\big[(\alpha+5)d(d+1)+1\big]} > 0, \tag{A.64}$$

where the inequality uses $\alpha \in [0,1]$ and $d > 0$. Hence $\widetilde{\pi}_b$ is a strictly concave quadratic. Substituting the optimal target prices (from Appendix A.6, equation (A.117)) gives

$$L = \frac{(5d^2+5d+1) - \alpha(5d^2+d)}{2(1-\alpha)d^2 + 4d + 1}, \tag{A.65}$$

and therefore

$$\frac{L}{2\kappa} = \frac{2\big[(\alpha+5)d(d+1)+1\big]}{(2d+1)\big[2(1-\alpha)d^2+4d+1\big]} = p_b^\dagger. \tag{A.66}$$

Completing the square in equation (A.62) yields

$$\widetilde{\pi}_b(p_b) = \widetilde{\pi}_b(p_b^\dagger) - \kappa(p_b - p_b^\dagger)^2, \qquad \widetilde{\pi}_b'(p_b) = -2\kappa(p_b - p_b^\dagger). \tag{A.67}$$

This establishes the quadratic claim and verifies directly that $p_b^\dagger$ is its unique maximizer. □

Proof of Proposition 10

Proof. To simplify the proof, we allow demand to be negative, so that the linear demand formula can be applied to the entire learning path.

By Lemma 2, firm b's perceived profit has the representation in equation (A.67) with $\kappa > 0$. Let $r_t \equiv r_{b,t}$, $e_t \equiv r_t - p_b^\dagger$, and $\mu \equiv \lambda\kappa$. Because perceived profit is quadratic, the divided difference equals its derivative at the midpoint:

$$\frac{\widetilde{\pi}_b(r_t) - \widetilde{\pi}_b(r_{t-1})}{r_t - r_{t-1}} = \widetilde{\pi}_b'\left(\frac{r_t + r_{t-1}}{2}\right) = -\kappa(e_t + e_{t-1}). \tag{A.68}$$

Whenever the two most recent prices yield unequal profits, equations (27)–(28) imply

$$e_{t+1} = (1-\mu)e_t - \mu e_{t-1}. \tag{A.69}$$

The characteristic polynomial is $z^2 - (1-\mu)z + \mu$. Its roots lie strictly inside the unit circle whenever $0 < \mu < 1$. Moreover,

$$\kappa = \frac{1}{4} + \frac{d}{2} - \frac{\alpha d}{2}\phi \le \frac{1+2d}{4}, \tag{A.70}$$

so $0 < \lambda < 4/(1+2d)$ implies $0 < \mu < 1$.

The same bound keeps the path in the relevant price interval. Under an ordinary update,

$$|e_{t+1}| \le (1-\mu)|e_t| + \mu|e_{t-1}| \le \max\{|e_t|, |e_{t-1}|\}. \tag{A.71}$$

Initial reference prices in $[0, 2p_b^\dagger]$ therefore imply $r_t \in [0, 2p_b^\dagger]$ for every t, ensuring that prices do not go negative. If two distinct reference prices yield the same profit, equation (A.67) implies that they are symmetric around $p_b^\dagger$; the midpoint selection convention selects $p_b^\dagger$. Equal-profit midpoint updates do not affect convergence: it can be verified that the two-period error maximum is nonincreasing between such events and, following the next ordinary update, is reduced by the factor $\mu < 1$. Thus, the conclusion holds whether midpoint events occur finitely or infinitely often. Hence $r_{b,t} \to p_b^\dagger$. Because finite-difference learning uses $p_{b,t} = r_{b,t}$, the realized price converges as well. □

Proof of Proposition 11

Proof. First, we establish the welfare benchmark in (i). When $\kappa_{CS} = \kappa_a = \kappa_b \equiv \kappa > 0$, the platform objective (32) is proportional to per-period total surplus,

$$TS^*(\alpha, \gamma) \;\equiv\; CS^*(\alpha, \gamma) + \pi_a^*(\alpha, \gamma) + \pi_b^*(\alpha, \gamma). \tag{A.72}$$

Under the linear demand system (19) (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

$$TS(p_a, p_b) \;=\; 4 - \frac{1}{8}\big(p_a^2 + p_b^2\big) - \frac{d}{4}\big(p_a - p_b\big)^2. \tag{A.73}$$

Hence $TS(p_a, p_b)$ is strictly decreasing in the level term $\frac{1}{8}(p_a^2 + p_b^2)$ and the dispersion term $\frac{d}{4}(p_a - p_b)^2$. Let $(p_a^*(\alpha, \gamma), p_b^*(\alpha, \gamma))$ denote the coercive equilibrium target prices induced by (α, γ). Proposition 6 states that both equilibrium prices are increasing in α and γ. Therefore, the smallest equilibrium prices occur at $(\alpha, \gamma) = (0, 0)$, which minimizes the level term. The dispersion term equals zero at $(0, 0)$ and is nonnegative everywhere, so it too is weakly smaller at $(0, 0)$ than at any (α, γ). Both terms enter (A.73) negatively, so

$$TS^*(\alpha, \gamma) \;\le\; TS^*(0, 0) \qquad \text{for all } (\alpha, \gamma) \in [0, 1]^2, \tag{A.74}$$

and $(\alpha, \gamma) = (0, 0)$ attains the maximum of the platform objective (32).

Next, we demonstrate (ii). Let the platform payoff be

$$\Omega(\alpha, \gamma) \equiv \kappa_{CS} CS^*(\alpha, \gamma) + \kappa_a \pi_a^*(\alpha, \gamma) + \kappa_b \pi_b^*(\alpha, \gamma), \tag{A.75}$$

and consider the difference between the maximal technology and the no-technology benchmark, $\widehat{\Delta} \equiv \Omega(1, 1) - \Omega(0, 0)$. Expanding,

$$\widehat{\Delta} = \kappa_{CS}\big(CS^*(1, 1) - CS^*(0, 0)\big) + \kappa_a\big(\pi_a^*(1, 1) - \pi_a^*(0, 0)\big) + \kappa_b\big(\pi_b^*(1, 1) - \pi_b^*(0, 0)\big). \tag{A.76}$$

Both sellers earn strictly more at $(1, 1)$ than at $(0, 0)$. For seller a, $\pi_a^*(0, 0)$ corresponds to the Bertrand profit. For linear demand, the condition in Proposition 5 holds, so $\pi_a^*(1, 1)$ strictly exceeds the symmetric collusive profit, which is strictly higher than the Bertrand profit. Thus, $\pi_a^*(1, 1) - \pi_a^*(0, 0) > 0$. For seller b, its incentive constraint binds at the selected incentive-compatible solution (equation (A.30)), so $\pi_b^*(1, 1)$ equals its best deviation payoff at $\alpha = 1$, which is its sequential-leader payoff $\max_q \pi_b\big(q, R_a(q)\big)$. Under the linear demand assumption, this exceeds the Bertrand payoff: $\pi_b^*(1, 1) - \pi_b^*(0, 0) > 0$.

Consumer surplus is strictly lower at $(1, 1)$: both equilibrium prices strictly exceed the Bertrand prices by Proposition 6 and consumer surplus is lower by Proposition 7. Therefore,

$CS^*(1,1) < CS^*(0,0)$. Define the threshold

$$\bar{\kappa}_{CS}(\kappa_a, \kappa_b) \equiv \frac{\kappa_a\big(\pi_a^*(1,1) - \pi_a^*(0,0)\big) + \kappa_b\big(\pi_b^*(1,1) - \pi_b^*(0,0)\big)}{CS^*(0,0) - CS^*(1,1)} \; > \; 0 \tag{A.77}$$

for any (κ_a, κ_b) with $\kappa_a + \kappa_b > 0$. Then for any $\kappa_{CS} \in [0, \bar{\kappa}_{CS})$, we have $\widehat{\Delta} > 0$, so the platform strictly prefers $(1,1)$ to $(0,0)$. Therefore $(0,0)$ cannot be optimal and any optimal policy satisfies $(\alpha^*, \gamma^*) \neq (0,0)$.

Finally, we establish (iii). Dividing the platform objective by κ_a gives

$$\frac{\Omega(\alpha,\gamma)}{\kappa_a} = \pi_a^*(\alpha,\gamma) + \frac{\kappa_{CS}}{\kappa_a} CS^*(\alpha,\gamma) + \frac{\kappa_b}{\kappa_a}\pi_b^*(\alpha,\gamma). \tag{A.78}$$

As κ_a grows with (κ_{CS}, κ_b) fixed, the last two terms vanish uniformly. Hence the platform's normalized objective converges uniformly to $\pi_a^*(\alpha,\gamma)$. Under linear demand, Lemma 3 establishes that π_a^* is uniquely maximized at the highest-speed, full-commitment policy. It follows that the platform's optimal choice of (α,γ) approaches $(1,1)$.

□

A.2 Alternative Timing Assumptions

This appendix examines whether the coercion mechanism depends on the move order at reprogramming dates. In the baseline model, 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. We consider four alternative timing structures for the reprogramming period.

Throughout this appendix, we maintain the baseline primitives and notation. Static profits are $\pi_i(p_i, p_j)$, the within-period speed advantage is $\alpha \in [0, 1]$, firm a's discount factor is $\beta \in [0, 1)$, and $\gamma \in [0, 1]$ governs algorithm persistence. We restrict attention to the same trigger algorithms used in the baseline analysis: firm a chooses an initialization price ρ and a target price $p_b^\dagger$ for firm b, and the update rule is

$$\sigma(p_b) = \begin{cases} \rho & \text{if } p_b = p_b^\dagger, \\ R_a(p_b) & \text{if } p_b \neq p_b^\dagger, \end{cases} \tag{A.79}$$

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

The alternatives below pertain to reprogramming-period timing. The first two correspond to different types of implementation delays, while the last two allow components of firm a's algorithm to be chosen before firm b moves. For all alternatives, in any non-reprogramming period, when a previously chosen algorithm remains in force, firm b faces the same incentive compatibility condition as in the baseline model:

$$\pi_b(p_b^\dagger, p_a^\dagger) \geq (1-\alpha)\pi_b(\hat{p}_b, p_a^\dagger) + \alpha\pi_b\big(\hat{p}_b, R_a(\hat{p}_b)\big) \quad \forall \hat{p}_b \neq p_b^\dagger. \tag{A.80}$$

Model (a): Forced 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 period $t = \tau$. However, here we assume that firm a employs its punishment rule $R_a(\cdot)$ after an update in period τ, regardless of whether firm b complies. In periods $t > \tau$ in which the algorithm persists, the algorithm operates as in the baseline model. This may represent a setting where the firm cannot credibly commit to leave its target price unchanged during a reprogramming period, or a setting where the firm might need to first demonstrate the punishment phase of its algorithm before it can successfully induce compliance.

Given initialization price ρ_τ for the first $1-\alpha$ fraction of the reprogramming period and the forced punishment after an update, firm b's reprogramming-period best response solves

$$p_{b\tau}^* \in \arg\max_{p_b} \left[(1-\alpha)\pi_b(p_b, \rho_\tau) + \alpha\pi_b\big(p_b, R_a(p_b)\big)\right]. \tag{A.81}$$

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. Unlike our baseline model, the equilibrium prices may

change from the reprogramming period to subsequent periods when the algorithm is in force. 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 response in this period.

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 period $t = \tau$. Here, assume that the within-period update technology $\sigma(\cdot)$ is not operative during the reprogramming period and becomes active only in future periods $t > \tau$ in which the algorithm persists. Thus firm a's initialization price determines its price for the entire reprogramming period. This timing could represent an implementation delay, where, for example, the firm needs to set up monitoring of its rival's price before implementing punishment. It is the converse of the reprogramming-period forced punishment in Model (a).

Maintaining the convention that firm a initializes at the target price, $\rho = p_a^\dagger$, firm b's reprogramming-period best response satisfies

$$p_{b\tau}^* \in \arg\max_{p_b} \pi_b(p_b, p_a^\dagger) = R_b(p_a^\dagger). \tag{A.82}$$

In later periods, once the update rule is active, the continuation target must satisfy (A.80). As in Model (a), prices in the reprogramming period may differ from the continuation target.

Model (c): Target Price Set First, then Simultaneous Price Setting. Now suppose firm a first commits only to the target price $p_b^\dagger$ for firm b. Immediately afterward, firm a chooses its initialization price ρ and firm b chooses p_b simultaneously. This timing captures settings in which the compliance target is more persistent or more easily communicated than actual prices.

Given a target $p_b^\dagger$, the reprogramming-period pricing subgame is

$$\rho^* \in \arg\max_{\rho | p_b^\dagger} \left[(1-\alpha)\pi_a(\rho, p_{b\tau}^*) + \alpha\pi_a\big(\sigma(p_{b\tau}^*), p_{b\tau}^*\big) + \frac{\beta\gamma}{1-\beta\gamma} V_c(\rho, p_b^\dagger; \alpha) \right], \tag{A.83}$$

$$p_{b\tau}^* \in \arg\max_{p_b | p_b^\dagger} \left[(1-\alpha)\pi_b(p_b, \rho^*) + \alpha\pi_b\big(p_b, \sigma(p_b)\big) \right], \tag{A.84}$$

where $V_c(\rho, p_b^\dagger; \alpha) = \pi_a(\rho, p_b^\dagger)$ if $\pi_b(p_b^\dagger, \rho) \geq D(\rho; \alpha)$. Otherwise, $V_c(\rho, p_b^\dagger; \alpha) = (1-\alpha)\pi_a(\rho, \hat{p}_b(\rho; \alpha)) + \alpha\pi_a\big(R_a(\hat{p}_b(\rho; \alpha)), \hat{p}_b(\rho; \alpha)\big)$ with $\hat{p}_b(\rho; \alpha)$ representing an optimal deviation by firm b, following (A.80). Using backward induction, firm a chooses the target $p_b^\dagger$ that yields the highest discounted payoff subject to firm b's best response in this subgame.

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): Forced Punishment 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	a as Follower	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 in 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 $(\alpha, \beta\gamma)$. The case $(0, 0)$ has neither speed nor commitment; $(0, 1)$ has full commitment but no speed advantage; $(1, 0)$ has maximum speed but no cross-period commitment; and $(1, 1)$ combines maximum speed and full commitment. Each entry characterizes target prices $(p_a^\dagger, p_b^\dagger)$; for the cases in which reprogramming-period prices differ (models (a) and (b) for $\beta\gamma = 1$), these correspond to the prices implemented in subsequent periods in which the algorithm remains in force.

Model (d): Algorithmic Stackelberg Leadership. Finally, suppose that at a reprogramming date firm a chooses the full algorithm first, including (ρ, σ) and the target $p_b^\dagger$. Firm b observes the algorithm and then chooses $p_{b\tau}$. In this sequential reprogramming-period timing, firm b solves

$$p_{b\tau}^* \in \arg \max_{p_b | (\rho, p_b^\dagger)} \left[(1-\alpha)\pi_b(p_b, \rho) + \alpha \pi_b\big(p_b, \sigma(p_b)\big) \right]. \tag{A.85}$$

If the target satisfies (A.80), then $p_{b\tau}^* = p_b^\dagger$ is optimal; otherwise firm b deviates immediately. Firm a's objective differs from the baseline because it can fully internalize firm b's response, even in the reprogramming period.

Table A-1 summarizes equilibria of the baseline timing and the four alternatives at the limiting 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 (14). The table reports the implemented target prices. For the cases in which prices in the initial reprogramming period differ from the target prices, the prices correspond to periods in which the algorithm persists.[32]

[32]For $\beta\gamma = 1$, the current-period payoff in a reprogramming period vanishes in firm a's objective function. For $\beta\gamma = 0$, target prices do not differ from reprogramming-period prices.

The timing variants yield two conclusions. First, maximal coercion is robust to the details of the reprogramming-period move order: across all variants, the same maximal-coercion target remains an equilibrium prediction under $(\alpha, \beta\gamma) = (1, 1)$. Second, the comparative statics continue to be organized by two forces: an enforcement technology, captured by speed α, and a source of commitment. Models (a) and (c) deliver the same limiting benchmarks as our baseline model. When the reprogramming period temporarily disables punishment, as in Model (b), speed has no impact without persistence, and the Bertrand outcome is obtained at $(\alpha, \beta\gamma) = (1, 0)$. In the baseline model and Models (a)–(c), persistence through $\beta\gamma$ supplies the commitment needed for maximal coercion. Model (d), by contrast, provides an additional source of commitment because firm a chooses the full algorithm before firm b sets its price. Thus, in Model (d), maximal coercion can be achieved when $\beta\gamma = 0$.

Existence and uniqueness under these timing assumptions require analogous model-specific single-peakedness and best-response slope conditions to those imposed by Appendix Assumption R1(i)–(ii) for the baseline problem. None of the variants requires Appendix Assumption R1(iii). The timing in Model (c) provides an intuitive parallel model: in our baseline, Appendix Assumption R1(iii) imposes that firm a does not consider a small unilateral change in $p_b^\dagger$ to affect whether p_{bt} is compliant. In Model (c), this happens through the move order, as firm a determines $p_b^\dagger$ before firm b sets its price, and the optimal choice of $p_b^\dagger$ can be determined by backward induction.

These alternative models clarify that the coercive mechanism is not a knife-edge consequence of a particular timing convention, but a result of the interaction of within-period enforcement and commitment.

A.3 Complementarity between Speed and Commitment

This appendix expands on the diagonal complementarity measure introduced in Section 4.1. When comparing values across the product-differentiation parameter, write firm a's equilibrium per-period profit as $\Pi_a(\alpha, \delta; d)$, where $\delta \equiv \beta\gamma$. For a common technology level $\xi = \alpha = \delta$, define

$$\Delta(\xi; d) \equiv \Pi_a(\xi, \xi; d) - \Pi_a(\xi, 0; d) - \Pi_a(0, \xi; d) + \Pi_a(0, 0; d). \tag{A.86}$$

This is the incremental gain from combining speed and commitment after subtracting the gain from speed alone and the gain from commitment alone. Equivalently, $\Delta(\xi; d) > 0$ means that adding speed is more valuable when the algorithmic firm also has commitment, or that adding commitment is more valuable when the algorithmic firm also has speed.

A related scale-free measure is

$$\rho(\xi, d) \equiv \frac{\Pi_a(\xi, \xi; d) - \Pi_a(0, \xi; d)}{\Pi_a(\xi, 0; d) - \Pi_a(0, 0; d)}. \tag{A.87}$$

The numerator is the value of adding speed when commitment is present at level ξ, and the denominator is the value of adding the same speed advantage with no commitment. Thus the same complementarity comparison in relative terms is given by $\rho(\xi, d) > 1$. The ratio is useful for comparisons when the gain from having only speed (the denominator) varies with d.

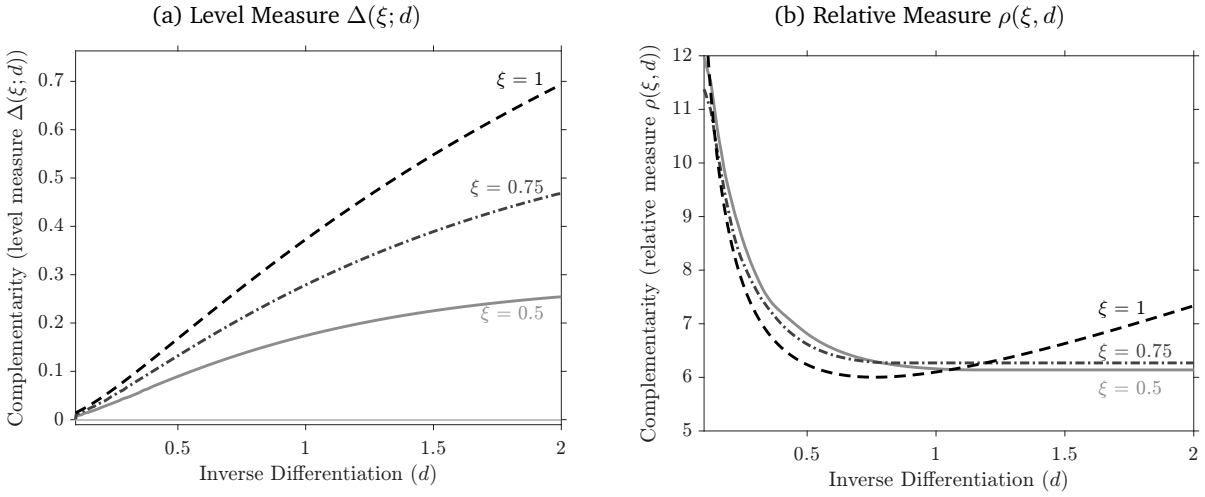
Figure A-1 plots the complementarity as a function of (inverse) product differentiation for $\xi \in \{0.5, 0.75, 1\}$. For this example, the complementarity is larger at higher values of ξ. The level measure, $\Delta(\xi; d)$, is increasing as products become closer substitutes. When products are highly differentiated, a rival's price has less effect on the algorithmic firm's demand, so the additional discipline created by combining speed and commitment is small. However, the ratio measure, $\rho(\xi, d)$, is largest when products are highly differentiated, because the gain from speed without commitment is small when competition between the firms is low. As substitutability rises, speed more strongly relaxes the rival's incentive constraint, increasing both the gains from speed and the complementarity of pairing speed and commitment. Panel (b) of Figure A-1 shows that the relative gain is non-monotonic in d.

For the linear demand system in equation (19), complementarity holds for any interior technology level and any degree of substitutability; at $\xi = 0$, the level measure equals zero by definition. We provide the analytical verification below.

Analytical verification Fix $d > 0$ and $\xi \in (0, 1)$. Given the maintained assumption about demand, write

$$A \equiv \frac{1 + 2d}{4}, \qquad B \equiv \frac{d}{2}, \qquad R(p) \equiv R_a(p) = R_b(p) = \frac{1 + Bp}{2A}. \tag{A.88}$$

Figure A-1: Complementarity between Speed and Commitment



Notes: Panel (a) reports the complementarity level $\Delta(\xi; d)$, the profit gain from combining speed and commitment at $\alpha = \beta\gamma = \xi$ net of the gains from each technology in isolation. Panel (b) reports $\rho(\xi, d)$, the value of adding speed when commitment is present relative to the value of adding speed when commitment is absent. Values of $\Delta(\xi; d) > 0$ and $\rho(\xi, d) > 1$ indicate complementarity. Demand follows equation (19) with zero marginal costs.

We begin with the two one-technology corners, $(\alpha, \delta) = (0, \xi)$ and $(\alpha, \delta) = (\xi, 0)$, where only one technology is active. When $\alpha = 0$ and $\delta = \xi$, firm b charges its static best response, and the commitment-only equilibrium is $(p_\xi, R(p_\xi))$. When $\alpha = \xi$ and $\delta = 0$, firm a is on its static best-response function, and the speed-only equilibrium is $(R(p_\xi), p_\xi)$. Solving the corresponding first-order and incentive-compatibility conditions gives the common one-technology corner price

$$p_\xi \equiv \frac{2(1+3d)}{1+4d+(3-\xi)d^2}. \tag{A.89}$$

The corner price pairs are $(p_\xi, R(p_\xi))$ and $(R(p_\xi), p_\xi)$. It suffices that

$$\begin{aligned} &\text{(C1)} \quad \bar{p}_b(\cdot; \xi) \text{ is weakly increasing and } p_a^\dagger(\xi, \xi; d) \geq p_\xi; \\ &\text{(C2)} \quad p_\xi \text{ is feasible at speed } \xi \text{ and } \bar{p}_b(p_\xi; \xi) - R_b(p_\xi) > 0; \\ &\text{(C3)} \quad 0 < \Pi_a(\xi, 0; d) - \Pi_a(0, 0; d) < Bp_\xi\big(\bar{p}_b(p_\xi; \xi) - R_b(p_\xi)\big). \end{aligned} \tag{A.90}$$

Let $(p_a^\dagger, p_b^\dagger)$ be the equilibrium at $(\alpha, \delta) = (\xi, \xi)$, so $p_b^\dagger = \bar{p}_b(p_a^\dagger; \xi)$. Under (C1)–(C2), $\bar{p}_b(p_\xi; \xi) \leq p_b^\dagger$ and p_ξ is feasible.

Write $\omega_\xi = \xi/(1-\xi)$. Regardless of whether firm b complies with the target $p_b^\dagger$, the current-period payoff for the target with $p_a = p_\xi$ is at least $\pi_a(p_\xi, p_b^\dagger)$. It equals this value under compliance, and static best-response punishment weakly raises it otherwise. Optimality and

the fact that products are substitutes therefore imply

$$(1+\omega_\xi)\Pi_a(\xi,\xi;d) \geq \pi_a\big(p_\xi, p_b^\dagger\big) + \omega_\xi \pi_a\big(p_\xi, \bar{p}_b(p_\xi;\xi)\big) \tag{A.91}$$

$$\geq (1+\omega_\xi)\pi_a\big(p_\xi, \bar{p}_b(p_\xi;\xi)\big). \tag{A.92}$$

Using the commitment-only corner and linearity in the rival's price, we have

$$\Pi_a(\xi,\xi;d) - \Pi_a(0,\xi;d) \geq Bp_\xi\big(\bar{p}_b(p_\xi;\xi) - R_b(p_\xi)\big). \tag{A.93}$$

Therefore,

$$\Delta(\xi;d) \geq Bp_\xi\big(\bar{p}_b(p_\xi;\xi) - R_b(p_\xi)\big) - \big(\Pi_a(\xi,0;d) - \Pi_a(0,0;d)\big) > 0, \tag{A.94}$$

where the last inequality is (C3). For linear demand, (C1) follows from Proposition 6 and the monotonicity of the upper incentive-compatible frontier, which can be verified directly, because $p_a^\dagger(\xi,\xi;d) \geq p_a^\dagger(0,\xi;d) = p_\xi$. Substituting p_ξ into the upper root of the quadratic incentive constraint (A.30) (from the proof of Proposition 6) verifies (C2), and substituting the two one-technology corner solutions shows that the speed-only profit gain is positive. Since $Bp_\xi\big(\bar{p}_b(p_\xi;\xi) - R_b(p_\xi)\big) > 0$, the remaining part of (C3) can be verified after squaring and clearing positive denominators; the resulting polynomial has positive coefficients for $\xi \in [0,1)$. Thus (C3) holds.

Demonstrating (C1), (C2), and (C3) yields $\Delta(\xi;d) > 0$ for $\xi \in (0,1)$. For $\xi = 1$, use the fact that, for the maintained linear demand system, the right-hand side of equation (A.94) converges to a strictly positive value as $\xi \to 1$. Therefore, by continuity, $\Delta(1;d) > 0$ for every d.

A.4 Private Incentives to Preserve a Speed Gap

This appendix isolates the speed margin in the linear-demand example with $d = 1$, zero marginal costs, and best-response punishment. It establishes two results. First, for each commitment level, firm b prefers to preserve a *speed gap*: whatever speed advantage and commitment level firm a has installed, firm b's profit-maximizing response leaves firm a with a strictly positive speed advantage. Second, when firm a's technology choice is endogenized and firm b can respond, backward induction delivers full commitment together with an interior speed advantage.

We first show that firm b prefers to preserve a speed gap for any technology installed by firm a. Write $\delta \equiv \beta\gamma$, and let $\Pi_i(\alpha, \delta)$ denote firm i's equilibrium per-period profit when α is firm a's speed advantage and δ is firm a's effective commitment. As elsewhere, $\delta = 1$ denotes the full-commitment limit. Fix the commitment level δ and let

$$\alpha_b^\dagger(\delta) \in \arg\max_{\alpha \in [0,1]} \Pi_b(\alpha, \delta) \tag{A.95}$$

denote firm b's preferred speed advantage. In our linear-demand example, $\Pi_b(\alpha, \delta)$ is single-peaked in α in equilibrium, and $\alpha_b^\dagger(\delta) > 0$ for all $\delta \in [0, 1]$, which we verify numerically. Thus, in this example, firm b's unconstrained optimum never eliminates the speed gap, whatever commitment level firm a has installed.[33] Table A-2 reports values for $\alpha_b^\dagger(\delta)$, together with both firms' induced profits and, for comparison, the profits when the speed gap is maximal, $\Pi_i(1, \delta)$. With a maximal speed gap, firm b's incentive constraint binds at its deviation payoff, which does not depend on δ when $\alpha = 1$, so $\Pi_b(1, \delta) = 0.762$ for every δ.

Take the full-commitment case as an example. At the two endpoints of the speed gap, firm b earns

$$\Pi_b(0, 1) = 0.823, \qquad \Pi_b(1, 1) = 0.762. \tag{A.96}$$

With no speed gap, firm a is the Stackelberg leader and firm b earns the sequential-follower payoff; with the maximal speed gap, maximal coercion holds firm b to its best deviation payoff. Firm b's preferred speed gap is interior:

$$\alpha_b^\dagger(1) = 0.328, \qquad \Pi_b(\alpha_b^\dagger(1), 1) = 0.951. \tag{A.97}$$

Thus, for $d = 1$, firm b prefers an interior speed gap both to no speed gap and to maximal coercion: a modest asymmetry softens price competition and lets firm b share in the additional surplus, while a maximal gap lets firm a extract too much.[34]

We next endogenize firm a's technology choice and show that this incentive persists when

[33] At $\delta = 0$, the preferred speed gap is maximal rather than interior: without commitment, speed makes firm a the Stackelberg follower, and firm b earns its sequential-leader payoff at $\alpha = 1$. For $\delta > 0$, $\alpha_b^\dagger(\delta)$ is interior.

[34] The optimal speed gap changes with d, but the qualitative preference for a speed gap remains.

Table A-2: Firm b's Preferred Speed Gap and Induced Profits

δ	$\alpha_b^\dagger(\delta)$	$\Pi_a(\alpha_b^\dagger(\delta),\delta)$	$\Pi_b(\alpha_b^\dagger(\delta),\delta)$	$\Pi_a(1,\delta)$	$\Pi_b(1,\delta)$
0.0	1.000	0.823	0.762	0.823	0.762
0.2	0.371	0.871	0.804	1.017	0.762
0.4	0.341	0.899	0.838	1.097	0.762
0.6	0.330	0.920	0.872	1.155	0.762
0.8	0.327	0.935	0.910	1.193	0.762
1.0	0.328	0.941	0.951	1.208	0.762

Notes: The table reports firm b's preferred speed gap $\alpha_b^\dagger(\delta)$ and the associated equilibrium per-period profits of both firms, for selected values of firm a's effective commitment $\delta \equiv \beta\gamma$. The final two columns report the equilibrium profits when the speed gap is maximal, $\alpha = 1$; the entry at $\delta = 1$ corresponds to maximal coercion. The calculations assume $d = 1$ under the linear demand system in equation (19), zero marginal costs, and best-response punishment.

both firms' adoption incentives are taken into account. Consider a simple two-stage adoption game. Firm a first chooses an installed speed advantage $\bar{\alpha} \in [0,1]$ and an effective commitment level $\delta \in [0,1]$. Firm b then decides how much of the speed gap to close, choosing the speed advantage $\alpha \in [0,\bar{\alpha}]$. For simplicity, we assume that adoption of any technology level is costless for both firms; we return to costly adoption below. The second-stage problem is

$$\alpha_b(\bar{\alpha},\delta) \in \arg\max_{\alpha\in[0,\bar{\alpha}]} \Pi_b(\alpha,\delta), \tag{A.98}$$

and single-peakedness gives $\alpha_b(\bar{\alpha},\delta) = \min\{\bar{\alpha}, \alpha_b^\dagger(\delta)\}$. Anticipating this response, firm a solves

$$(\bar{\alpha}^*,\delta^*) \in \arg\max_{\bar{\alpha}\in[0,1],\,\delta\in[0,1]} \Pi_a(\alpha_b(\bar{\alpha},\delta),\delta)\,. \tag{A.99}$$

Because firm a's profit is increasing in the speed advantage for fixed δ, firm a installs enough speed to accommodate firm b's preferred residual gap, and the first-stage problem reduces to the one-dimensional choice $\delta^* \in \arg\max_{\delta\in[0,1]} \Pi_a(\alpha_b^\dagger(\delta),\delta)$.

Firm a's induced profit is increasing in δ and is maximized at the full-commitment endpoint, so backward induction yields $\delta^* = 1$. Firm b then leaves an interior speed gap, with $\alpha_b^\dagger(1) = 0.328$, $\Pi_a(\alpha_b^\dagger(1),1) = 0.941$, and $\Pi_b(\alpha_b^\dagger(1),1) = 0.951$. Firm a's payoff exceeds both its payoff when the speed gap is eliminated ($\Pi_a(0,1) = 0.762$) and its payoff when it lacks commitment ($\Pi_a(\alpha_b^\dagger(0),0) = 0.823$). Thus the adoption game delivers the same economic conclusion as the direct payoff comparison: firm a chooses full commitment, while firm b closes only part of the speed gap and leaves firm a with a speed advantage. If there are costs for firm a to install a greater initial speed advantage, it may choose $\bar{\alpha}^* = \alpha_b^\dagger(1) = 0.328$ together with $\delta^* = 1$ directly; with no costs, firm a is indifferent among installed speed advantages $\bar{\alpha} \geq \alpha_b^\dagger(1)$, because firm

b's second-stage choice leaves the same speed gap.

If firm b faces a convex cost of reducing the gap, the conclusion remains. Suppose it costs firm b $C_b(\bar{\alpha} - \alpha)$ to close the gap from $\bar{\alpha}$ to α. The second-stage objective becomes $\Pi_b(\alpha, \delta) - C_b(\bar{\alpha} - \alpha)$. Relative to the zero-cost problem, the cost term weakly shifts firm b's optimum toward a larger speed advantage. With sufficiently high marginal costs, firm b may optimally make no speed investment. Costs for firm a can reduce its desired installed speed or commitment, but do not alter the zero-cost result that equal access to technologies need not eliminate asymmetries in their use.

A.5 One Algorithmic Firm in an N-Firm Oligopoly

This appendix provides the formal details behind Section 4.3. There is one algorithmic firm, denoted by a, and $m = N - 1$ rivals that do not use algorithms. Let $\mathbf{r} = (r_1, \ldots, r_m)$ denote the vector of slow rivals' prices, and let $\mathbf{r}_{-j}$ denote the vector excluding rival j's price. We provide the derivation for best-response punishment. Under marginal-cost punishment, the derivation is identical except that firm a's price in the deviating rival's payoff is set equal to marginal cost rather than to its static best response, reducing the best deviation payoff and increasing equilibrium prices.

Firm a chooses a target price for itself, $p_a^\dagger$, and a vector of target prices for the slow rivals, $\mathbf{r}^\dagger = (r_1^\dagger, \ldots, r_m^\dagger)$. In this appendix, $\pi_j(p_a, r_j, \mathbf{r}_{-j})$ denotes slow rival j's profit when the algorithmic firm charges p_a, rival j charges r_j, and the other slow rivals charge $\mathbf{r}_{-j}$. The oligopoly analog of the target-price problem in the baseline model is

$$
\begin{aligned}
\max_{(p_a^\dagger, \mathbf{r}^\dagger)|\mathbf{r}_t} \quad & (1-\alpha)\pi_a(\rho, \mathbf{r}_t) + \alpha\pi_a(\sigma(\mathbf{r}_t), \mathbf{r}_t) + \frac{\beta\gamma}{1-\beta\gamma}\pi_a(p_a^\dagger, \mathbf{r}^\dagger) && \text{(A.100)} \\
\text{s.t.} \quad (i) \quad & \rho = p_a^\dagger, \\
(ii) \quad & \sigma(\mathbf{r}) = \begin{cases} p_a^\dagger, & \text{if } \mathbf{r} = \mathbf{r}^\dagger, \\ R_a(\mathbf{r}), & \text{if } \mathbf{r} \neq \mathbf{r}^\dagger, \end{cases} \\
(iii) \quad & \pi_j(p_a^\dagger, r_j^\dagger, \mathbf{r}_{-j}^\dagger) \geq (1-\alpha)\pi_j(p_a^\dagger, \widehat{r}_j, \mathbf{r}_{-j}^\dagger) + \alpha\pi_j\Big(R_a(\widehat{r}_j, \mathbf{r}_{-j}^\dagger), \widehat{r}_j, \mathbf{r}_{-j}^\dagger\Big) \\
& \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \forall \widehat{r}_j, \quad \forall j \in \{1, \ldots, m\}.
\end{aligned}
$$

The final inequality is the implementability condition for the target vector. Slow rival j compares its on-path payoff to the best payoff from a unilateral deviation, taking the other slow rivals' target prices as fixed. In contrast to the duopoly case, a deviating slow rival can also steal demand from other slow firms, whose prices do not adjust within the period.

We now specialize to the symmetric linear-demand system used in Section 4.3:

$$
D_i(p_i, p_{-i}) = 1 - \left(\frac{1}{4} + \frac{d}{2}\right) p_i + \frac{d}{2}\overline{p}_{-i}, \tag{A.101}
$$

where $\overline{p}_{-i}$ is the average price charged by firm i's rivals. The symmetric Bertrand and joint-profit-maximizing prices are, respectively, $p^B(d) = 2/(1+d)$, and $p^C = 2$.

We restrict attention to symmetric targets for the slow rivals. Let $\mathbf{1}_m$ be the m-vector of ones, write $\mathbf{r}^\dagger = r\mathbf{1}_m$, and write p for firm a's target price. On the target path, firm a's demand is

$$
D_a^m(p, r; d) = 1 - \left(\frac{1}{4} + \frac{d}{2}\right) p + \frac{d}{2} r \tag{A.102}
$$

and each slow rival's demand is

$$D_s^m(p,r;d) = 1 + \frac{d}{2m}p - \left(\frac{1}{4} + \frac{d}{2m}\right)r \tag{A.103}$$

Consider the maximal-coercion limit, $\alpha = 1$ and $\beta\gamma \to 1$, under best-response punishment. If one slow rival deviates to $\hat{r}$ while the other $m-1$ slow rivals remain at r, firm a's best response is

$$R_a(\hat{r}, r; m) = \frac{2 + \frac{d}{m}\big(\hat{r} + (m-1)r\big)}{1 + 2d}. \tag{A.104}$$

Substituting this response into the deviating rival's payoff gives

$$\hat{r}\left[L_m(r;d) - \kappa_m(d)\hat{r}\right], \tag{A.105}$$

where

$$\kappa_m(d) \equiv \frac{1}{4} + \frac{d}{2} - \frac{d^2}{2m^2(1+2d)} \tag{A.106}$$

and

$$L_m(r;d) \equiv \left(1 + \frac{d}{m(1+2d)}\right)\left(1 + \frac{d(m-1)}{2m}r\right). \tag{A.107}$$

The maximal deviation payoff is therefore $L_m(r;d)^2/[4\kappa_m(d)]$. The symmetric target (p,r) is incentive compatible if and only if

$$r\left[1 + \frac{d}{2m}p - \left(\frac{1}{4} + \frac{d}{2m}\right)r\right] \geq \frac{L_m(r;d)^2}{4\kappa_m(d)}. \tag{A.108}$$

At the coercive optimum this constraint binds. Otherwise, firm a could slightly raise the common slow-rival target price, increasing its own demand while preserving incentive compatibility. Solving the binding constraint for p gives

$$\widetilde{p}_m(r;d) = \frac{2m}{d}\left\{\frac{L_m(r;d)^2}{4\kappa_m(d)r} - 1 + \left(\frac{1}{4} + \frac{d}{2m}\right)r\right\}. \tag{A.109}$$

Hence maximal coercion under best-response punishment reduces to the problem

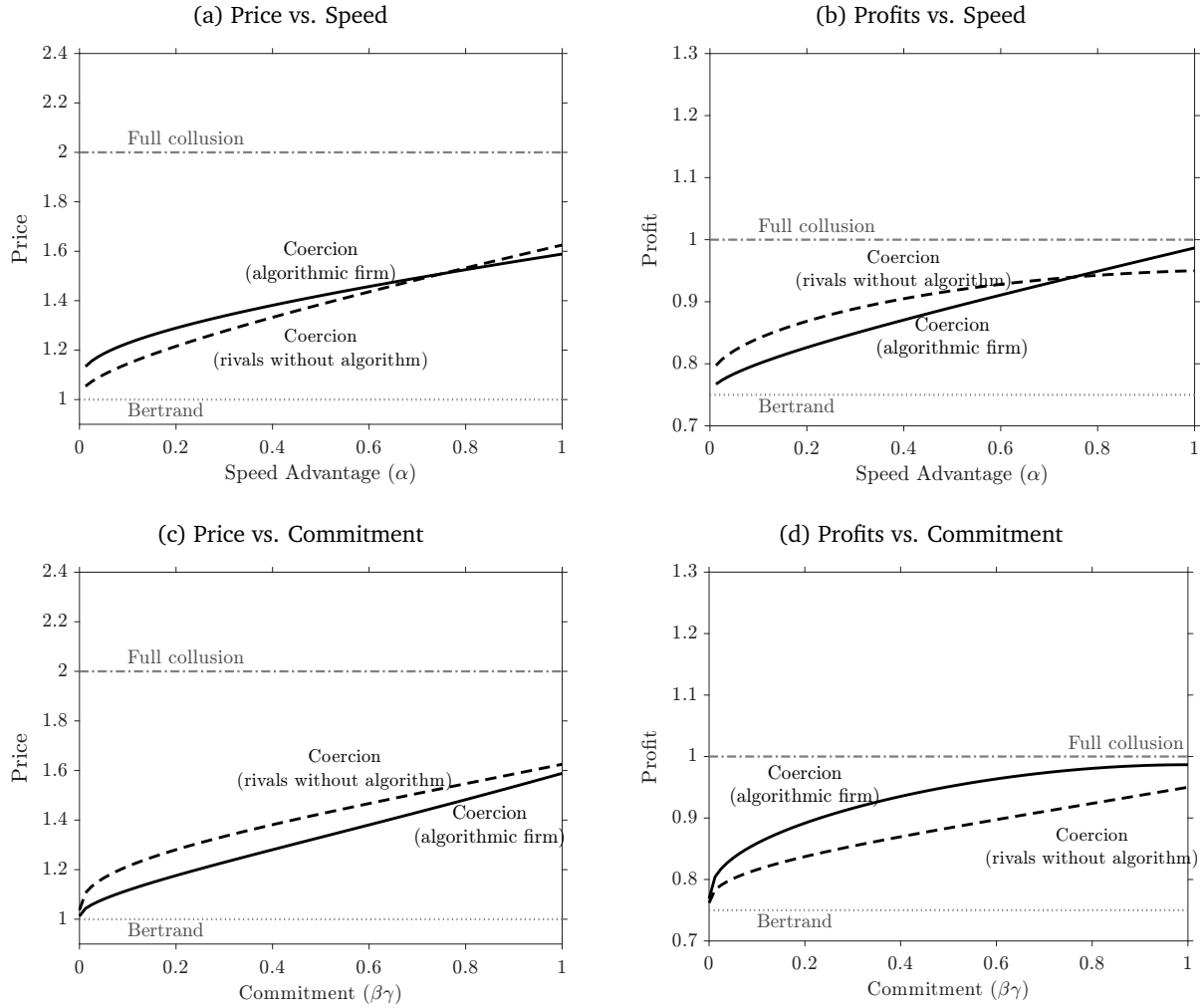
$$\max_{r \geq 0} \quad \widetilde{p}_m(r;d)\left[1 - \left(\frac{1}{4} + \frac{d}{2}\right)\widetilde{p}_m(r;d) + \frac{d}{2}r\right]. \tag{A.110}$$

Let $r_m^\star(d)$ denote the maximizer and define $p_m^\star(d) = \widetilde{p}_m(r_m^\star(d);d)$. The corresponding quantity-weighted average price is

$$\overline{p}_m(d) = \frac{p_m^\star(d)D_a^m(p_m^\star(d), r_m^\star(d);d) + mr_m^\star(d)D_s^m(p_m^\star(d), r_m^\star(d);d)}{D_a^m(p_m^\star(d), r_m^\star(d);d) + mD_s^m(p_m^\star(d), r_m^\star(d);d)}. \tag{A.111}$$

Qualitatively, speed and commitment have similar effects in oligopoly as in the duopoly

Figure A-2: Prices and Profits in Coercive Equilibrium with Three Firms

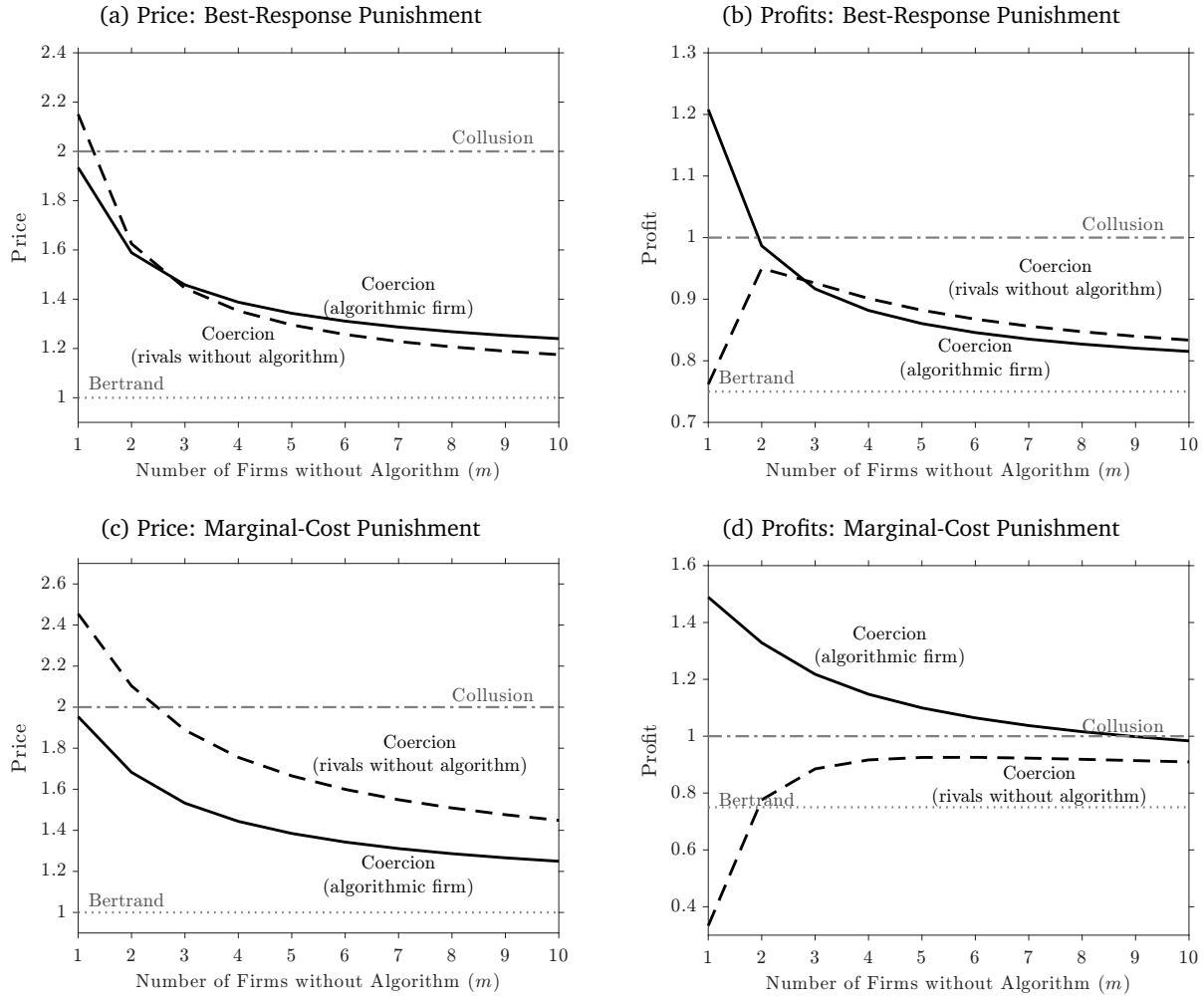


Notes: Panels (a) and (b) show equilibrium prices and profits for different values of speed advantage α under coercion with full commitment ($\beta\gamma = 1$). Panels (c) and (d) show equilibrium prices and profits under coercion for different values of commitment, $\beta\gamma$, when the faster firm has maximum speed ($\alpha = 1$). Outcomes under Bertrand competition with simultaneous pricing and joint profit maximization are displayed for comparison. Demand follows the linear specification in equation (A.101) with $d = 1$.

case. To illustrate, we simulate equilibrium outcomes for the $N = 3$ case with $d = 1$. Appendix Figure A-2 shows equilibrium prices (panel (a)) and profits (panel (b)) with full commitment and different values for the speed advantage. Panels (c) and (d) of Appendix Figure A-2 show equilibrium prices and profits in the case with maximum speed advantage and different values for the degree of commitment. As in Figures 3 and 4, prices and firm a's profits are increasing in its speed advantage and commitment. However, the ability of firm a to coerce rivals into setting higher prices is somewhat muted compared to the two-firm case due to additional slow rivals and the change in the incentive compatibility constraint.

Finally, we examine the maximal coercion equilibrium by the number of firms in Figure A-

Figure A-3: Prices and Profits in Maximal Coercion under Oligopoly



Notes: Panel (a) shows prices and panel (b) shows profits under Bertrand competition, maximal coercion with maximum speed and commitment, and joint-profit maximization with the baseline best-response punishment. Panel (c) shows prices and panel (d) shows profits under marginal-cost punishment. In panels (b) and (d), profits are per firm; the dashed coercion curve reports the profit of each symmetric rival without an algorithm. The horizontal axis reports the number of such rivals, $m = N - 1$, so the total number of firms is $N = m + 1$. Demand follows the linear specification in equation (A.101) with $d = 1$.

3. Panels (a) and (b) show the equilibrium prices and profits with best-response punishment, while panels (c) and (d) show the equilibrium prices and profits with (the harsher) marginal-cost punishment. For any value of d, the outcomes under Bertrand and collusion are constant as the number of firms increases. Coercive equilibrium prices are decreasing as the algorithmic firm faces additional rivals without an algorithm, indicating that increased competition can limit coercive outcomes. However, prices and profits are still substantially higher relative to the Bertrand equilibrium, even with several rivals. With ten rival firms, market prices are on average 18 percent higher than Bertrand prices under best-response punishment and 43 percent higher under marginal-cost punishment.

A.6 First-Order Conditions for the Linear Pricing Rule

This appendix presents the formal conditions behind Section 5.2. The algorithmic firm chooses a target price vector $(p_a^\dagger, p_b^\dagger)$ and uses a linear update rule $\sigma(p_b) = \phi p_b$ with $\phi = p_a^\dagger / p_b^\dagger$. Its target problem can be written as

$$\begin{aligned} &\max_{(p_a^\dagger, p_b^\dagger)} \pi_a(p_a^\dagger, p_b^\dagger) \\ &\text{s.t.} \quad p_b^\dagger = \operatorname*{argmax}_{p_b} (1-\alpha)\pi_b(p_b, p_a^\dagger) + \alpha \pi_b(p_b, \phi p_b). \end{aligned} \tag{A.112}$$

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

$$(1-\alpha)\frac{\partial \pi_b(p_b, p_a^\dagger)}{\partial p_b} + \alpha \left[\frac{\partial \pi_b(p_b, \phi p_b)}{\partial p_b} + \phi \, \frac{\partial \pi_b(p_b, \phi p_b)}{\partial p_a} \right] = 0. \tag{A.113}$$

Because $\phi = p_a^\dagger / p_b^\dagger$, the slower firm's first-order condition, evaluated at $p_b = p_b^\dagger$, defines an implementability relation between the target prices. Locally, write this relation as $p_a^\dagger = p_a^{\mathrm{I}}(p_b^\dagger)$, where $p_a^{\mathrm{I}}(p_b^\dagger)$ is the target price for firm a implied by the slower firm's first-order condition. The algorithmic firm therefore chooses $p_b^\dagger$ along this implementability relation to maximize $\pi_a(p_a^{\mathrm{I}}(p_b^\dagger), p_b^\dagger)$. Its first-order condition is

$$\frac{d}{dp_b^\dagger} \pi_a\big(p_a^{\mathrm{I}}(p_b^\dagger), p_b^\dagger\big) = \frac{\partial \pi_a\big(p_a^{\mathrm{I}}(p_b^\dagger), p_b^\dagger\big)}{\partial p_a} \, \frac{dp_a^{\mathrm{I}}(p_b^\dagger)}{dp_b^\dagger} + \frac{\partial \pi_a\big(p_a^{\mathrm{I}}(p_b^\dagger), p_b^\dagger\big)}{\partial p_b} = 0. \tag{A.114}$$

Together, these conditions characterize the steady-state target when the algorithmic firm is constrained to use a linear pricing rule.

Under the linear demand system (19) with zero marginal costs and the linear rule $\sigma(p_b) = \phi p_b$ with $\phi = p_a^\dagger / p_b^\dagger$, the two conditions above can be solved in closed form. Substituting $\pi_b(p_b, p_a) = p_b\big[1 - (\frac{1}{4} + \frac{d}{2})p_b + \frac{d}{2} p_a\big]$, firm b's within-period average profit is the quadratic

$$(1-\alpha)\pi_b(p_b, p_a^\dagger) + \alpha \, \pi_b(p_b, \phi p_b) = \Big[1 + \tfrac{d}{2}(1-\alpha) p_a^\dagger\Big] p_b - \Big[\tfrac{1}{4} + \tfrac{d}{2} - \tfrac{\alpha d}{2}\phi\Big] p_b^2. \tag{A.115}$$

Setting its derivative to zero at $p_b = p_b^\dagger$ and using $\phi \, p_b^\dagger = p_a^\dagger$ yields the implementability relation

$$(1 + 2d)\, p_b^\dagger \;=\; 2 + d(1+\alpha)\, p_a^\dagger. \tag{A.116}$$

Substituting (A.116) into $\pi_a(p_a^\dagger, p_b^\dagger)$ gives a strictly concave quadratic in $p_a^\dagger$ alone, with maxi-

mizer and implied values

$$
\begin{aligned}
p_a^\dagger &= \frac{2(3d+1)}{2(1-\alpha)d^2+4d+1}, \\
p_b^\dagger &= \frac{2\big[(\alpha+5)d(d+1)+1\big]}{(2d+1)\big[2(1-\alpha)d^2+4d+1\big]}, \\
\phi &= \frac{(2d+1)(3d+1)}{(\alpha+5)d(d+1)+1}.
\end{aligned}
\tag{A.117}
$$

At $\alpha = 0$ the rule provides no within-period response, and (A.117) reduces to the sequential-leader benchmark: $p_a^\dagger = 2(3d+1)/(2d^2+4d+1)$ with $p_b^\dagger = R_b(p_a^\dagger)$. Both targets are increasing in α.

A.7 Alternative Learning Rules

The main text uses finite-difference gradient learning as the leading case. This subsection shows that convergence is not specific to that procedure by considering three alternative methods that search the same perceived profit function using only firm b's own prices and realized profits. As in Section 5.1, let $r_{b,t}$ denote the current reference price, $\widehat{g}_{b,t}$ the adjustment signal, and λ_t the step size in equation (26). Let $\Pi_{b,t}$ denote realized profit in period t. All prices are constrained to be nonnegative. If a proposed price is negative, firm b sets it equal to zero. For the finite-difference benchmark in the simulation comparison, we use the main-text rule: $p_{b,t} = r_{b,t}$, $\widehat{g}_{b,t}$ is the finite-difference estimate in equation (27), and $\lambda_t = 0.4$. Its initial reference prices are $r_{b,0} = p_b^0$ and $r_{b,1} = p_b^0 + 0.05$.

Paired A/B Price Experiments. This approach asks whether a small price increase or a small price decrease from the current reference price is more profitable. Firm b implements this using pairs of successive periods. In each odd period, it randomly chooses whether to try a price below or above the reference price; in the following even period, it tries the other price. Let $\mathbb{1}_t$ denote an indicator equal to 1 if t is even and 0 if t is odd. Let $z_t \in \{-1, 1\}$ be chosen with equal probability when t is odd and set $z_t = -z_{t-1}$ when t is even. Realized prices are

$$p_{b,t} = r_{b,t} + z_t h_t, \qquad h_t > 0, \tag{A.118}$$

where $r_{b,t+1} = r_{b,t}$ when t is odd.

Once both prices in a pair have been observed—i.e., in even periods—firm b forms the finite-difference estimate

$$\widehat{g}_{b,t} = \frac{\Pi_{b,t} - \Pi_{b,t-1}}{p_{b,t} - p_{b,t-1}}. \tag{A.119}$$

The reference price is updated according to

$$r_{b,t+1} = r_{b,t} + \mathbb{1}_t \, \lambda_t \, \widehat{g}_{b,t}. \tag{A.120}$$

The indicator $\mathbb{1}_t$ makes explicit that the reference price is left unchanged after odd periods and is updated only after even periods, once $\widehat{g}_{b,t}$ is available.

Let $k \equiv \lfloor (t+1)/2 \rfloor$ denote the index of the round to which period t belongs, so that periods $2k-1$ and $2k$ together form round k. The simulations set $r_{b,1} = p_b^0$, $\lambda_t = 0.4$ whenever $\mathbb{1}_t = 1$, and $h_t = 0.01/k$, so that the experimental radius shrinks with the number of completed rounds. The arm order is randomized independently across simulation paths and rounds.

Quasi-Newton Learning. This method uses the firm's three most recent prices and profits. Firm b begins by choosing p_b^0, $p_b^0 + 0.05$, and $p_b^0 + 0.10$. As in the main text, for $t \geq 1$, define the

estimated profit slope as

$$\widehat{g}_{b,t} = \frac{\Pi_{b,t} - \Pi_{b,t-1}}{p_{b,t} - p_{b,t-1}}. \tag{A.121}$$

The change between two profit slopes, which reflects the curvature, can be estimated by

$$\widehat{h}_{b,t} = \frac{2\left(\widehat{g}_{b,t} - \widehat{g}_{b,t-1}\right)}{p_{b,t} - p_{b,t-2}}. \tag{A.122}$$

The reference price is updated using the midpoint of the prior two prices. Let $\lambda_t = -1/\widehat{h}_{b,t} > 0$. The reference price is updated according to

$$r_{b,t+1} = \frac{r_{b,t} + p_{b,t-1}}{2} + \lambda_t \, \widehat{g}_{b,t}, \tag{A.123}$$

so that the curvature determines the step size. We assume that the firm always chooses to set the reference price, $p_{b,t} = r_{b,t}$. If two consecutive prices coincide, firm b continues to choose that price.

Adaptive Directional Price Search. This procedure does not estimate the profit slope. Instead, firm b tests a price a step away from its current reference price in a fixed direction. If the price raises profit, the firm accepts the new price and continues experimenting in the same direction. If the price does not raise profit, the firm keeps the prior reference price, reverses the direction of the next experiment, and shrinks the step size.

Let $\widehat{g}_{b,t-1} \in \{-1, 1\}$ denote the search direction used in period t, and let $h_t > 0$ denote the experiment size. The initial price is the initial reference price in period $t = 0$, so that $p_{b,0} = r_{b,0}$. For $t \geq 1$ the realized price is

$$p_{b,t} = r_{b,t} + h_t \, \widehat{g}_{b,t-1}. \tag{A.124}$$

Let $\Pi_{b,t}$ denote realized profit at the price $p_{b,t}$, and let $\Pi^r_{b,t}$ denote realized profit the last time firm b selected the incumbent reference price $r_{b,t}$. The search direction is updated according to

$$\widehat{g}_{b,t} = \begin{cases} \widehat{g}_{b,t-1}, & \Pi_{b,t} > \Pi^r_{b,t}, \\ -\widehat{g}_{b,t-1}, & \text{otherwise.} \end{cases} \tag{A.125}$$

The experiment size declines after the realized price fails to increase profits and remains the same after a profit increase, $h_{t+1} = \chi h_t + \mathbb{1}[\Pi_{b,t} > \Pi^r_{b,t}](1 - \chi)h_t$ for $t \geq 1$, where $\chi \in (0, 1)$ is a contraction factor. Then, the reference price is updated according to

$$r_{b,t+1} = r_{b,t} + \lambda_t \, \widehat{g}_{b,t}, \tag{A.126}$$

where $\lambda_t = \mathbb{1}[\Pi_{b,t} > \Pi^r_{b,t}]h_t$, so that the reference price follows the experiment step if profits increase and remains fixed otherwise. The simulations set $r_{b,0} = p^0_b$, $r_{b,1} = r_{b,0}$, $\widehat{g}_{b,0} = 1$,

Table A-3: Periods to Convergence under Alternative Learning Rules

Learning Rule	Mean Periods to Convergence
Finite-difference gradient learning	20.2
Paired A/B price experiments	46.9
Quasi-Newton learning	4.0
Adaptive directional price search	25.4

Notes: The table reports mean convergence times across 1,000 initial prices under the calibration above. All four methods converge from every initial value, using the same realized-price convergence criterion. Each paired A/B experimental round counts as two periods, and the three initial quasi-Newton experiments count as periods.

$h_1 = 0.05$, and $\chi = 0.5$.

Simulation Comparison

In order to compare the approaches, we set $d = 1$ and $\alpha = 1$. Under the optimal linear rule given by equation (A.117), firm a updates following $p_a = (12/13)p_b$ and firm b faces the perceived profit function

$$\widetilde{\pi}_b(p_b) = p_b - \frac{15}{52}p_b^2. \tag{A.127}$$

This is uniquely maximized at $p_b^\dagger = 26/15$. The corresponding target for firm a is $p_a^\dagger = 8/5$. Each method is simulated from 1,000 evenly spaced initial reference prices $r_{b,0} = p_b^0$ with $p_b^0 \in [0, 2]$, and firm b observes only its own prices and realized profits.

For every method, convergence is defined by the first of ten consecutive periods in which the realized price is within 0.1 percent of $p_b^\dagger$. Each two-price A/B experimental round counts as two periods, and all initial price experiments count as periods.

The variation in convergence time reflects both how each rule uses profit feedback and how frequently it updates its price. Quasi-Newton learning is especially fast in this environment because perceived profit is exactly quadratic. The first three observations reveal exactly how the profit slope changes with price, so the first quasi-Newton update selects the profit-maximizing price. With nonquadratic profits, recent changes in the slope would provide only a local guide, and convergence would generally be slower. Paired A/B testing is slower because each update requires two price realizations and our convergence criterion requires the experimental radius to become small. Directional search falls between these cases.

A.8 Calendar-Time Mapping of Speed and Commitment

Table A-4: Calendar-Time Mapping of Speed and Commitment

Panel (a): Speed Advantage $\alpha(T_a, T_b)$

	Firm b Pricing Interval T_b				
Firm a pricing interval T_a	Quarterly	Monthly	Weekly	Daily	Hourly
Quarterly	0.000				
Monthly	0.667	0.000			
Weekly	0.917	0.750	0.000		
Daily	0.988	0.964	0.857	0.000	
Hourly	1.000	0.999	0.994	0.958	0.000

Panel (b): Commitment Parameter $\gamma(T_b)$

	Firm b Pricing Interval T_b				
Firm a pricing interval T_a	Quarterly	Monthly	Weekly	Daily	Hourly
Quarterly	0.500				
Monthly	0.500	0.794			
Weekly	0.500	0.794	0.944		
Daily	0.500	0.794	0.944	0.992	
Hourly	0.500	0.794	0.944	0.992	1.000

Notes: Panel (b) assumes Poisson arrivals of reprogramming opportunities and sets the probability that the algorithm remains unchanged over one quarter to 0.500. Blank entries correspond to cases in which firm b prices faster than firm a. Calculations use one quarter equal to three four-week months, one week equal to seven days, and one day equal to 24 hours. Entries are rounded to three decimal places.

We describe how benchmark pricing frequencies map into the speed advantage parameter α and the commitment parameter γ. Let T_a and T_b denote the time intervals between pricing opportunities for firms a and b. By definition, a period in the model has length T_b. We assume that pricing opportunities for firm b always coincide with a pricing opportunity for firm a. With $T_a \leq T_b$, firm a first responds after T_a units of time and its updated price applies for the remaining $T_b - T_a$ units. Hence $\alpha(T_a, T_b) = (T_b - T_a)/T_b = 1 - T_a/T_b$.

Panel (a) of Table A-4 presents example values of α for quarterly, monthly, weekly, daily, and hourly pricing. We use one quarter equal to three months, one month equal to four weeks, one week equal to seven days, and one day equal to 24 hours. Entries with $T_a > T_b$ are left blank because they correspond to b as the faster firm. With $T_a = T_b$, $\alpha = 0$, and as $T_a \to 0$, $\alpha \to 1$.

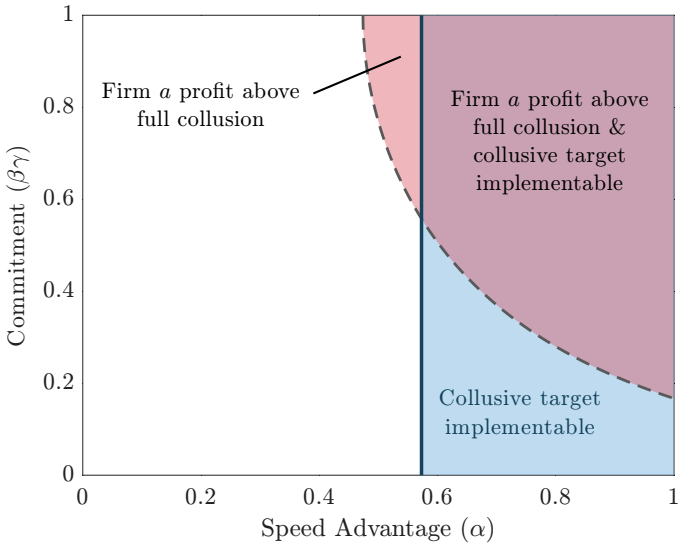
To map these frequencies to commitment, assume that reprogramming opportunities arrive according to a Poisson process such that at least one opportunity arrives with probability 0.500 each quarter. If an opportunity arrives during a period, the algorithmic firm can reprogram its algorithm at the start of the next period. The probability that the algorithm stays fixed over a period is then $\gamma(T_b) = 0.500^{\,T_b/T_Q}$, where T_Q denotes one quarter.

Panel (b) reports example values for γ. Holding the calendar-time reprogramming process

fixed, faster pricing by firm b both reduces firm a's speed advantage and raises algorithm persistence across periods. Note that γ depends on T_b but not on T_a. To map pricing frequencies into effective commitment $\beta\gamma$, γ must be multiplied by the implied discount factor β for firm b's pricing interval T_b. Annual discount factors are often 0.9 or higher, implying discount factors close to one for pricing intervals shorter than one year.

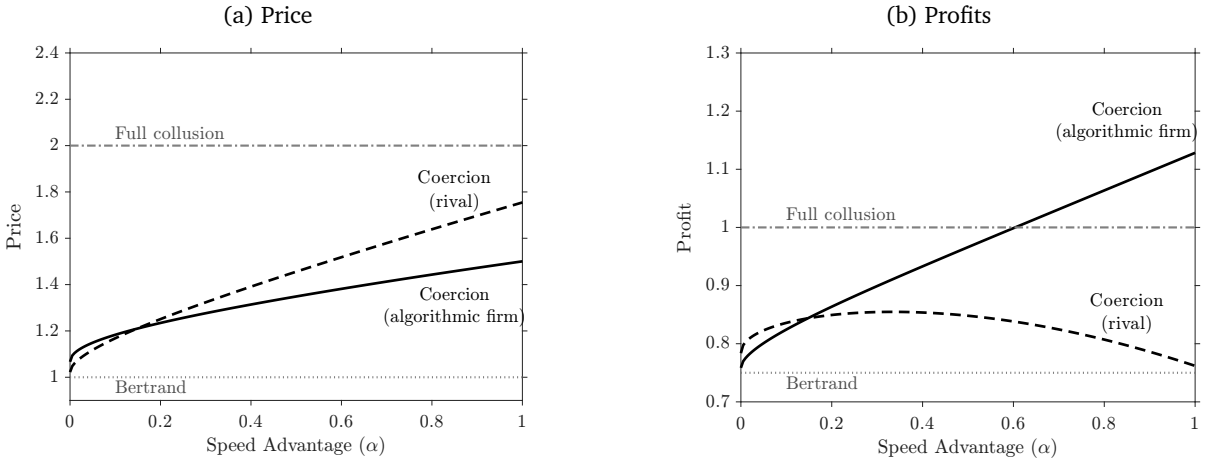
A.9 Additional Tables and Figures

Figure A-4: Implementability of Collusive Benchmark



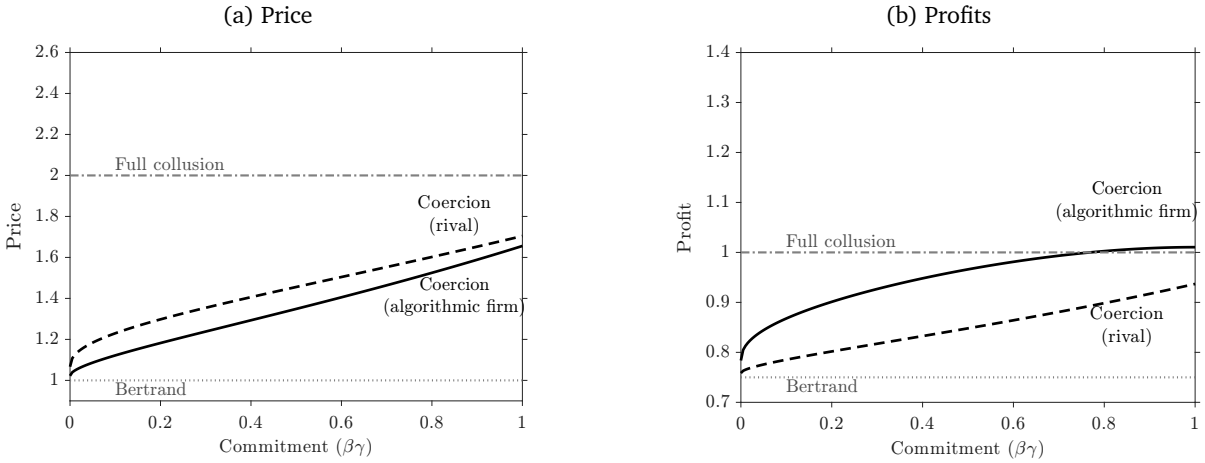
Notes: The figure plots the region where full collusion is implementable as a function of the algorithmic firm's speed advantage α and commitment $\beta\gamma$ in blue. For comparison, the dashed line provides the boundary of the region in which firm a earns more than its share of full collusion profits. Feasible implementation of collusive prices only depends on speed because the relevant condition is firm b's current-period deviation incentive; however, even with α sufficiently high, firm a may earn less than collusive profits in equilibrium when commitment is low, as illustrated in the lower right region of the figure. Demand follows the linear specification in equation (19) with $d = 1$.

Figure A-5: Prices and Profits in Coercive Equilibrium: Partial Commitment



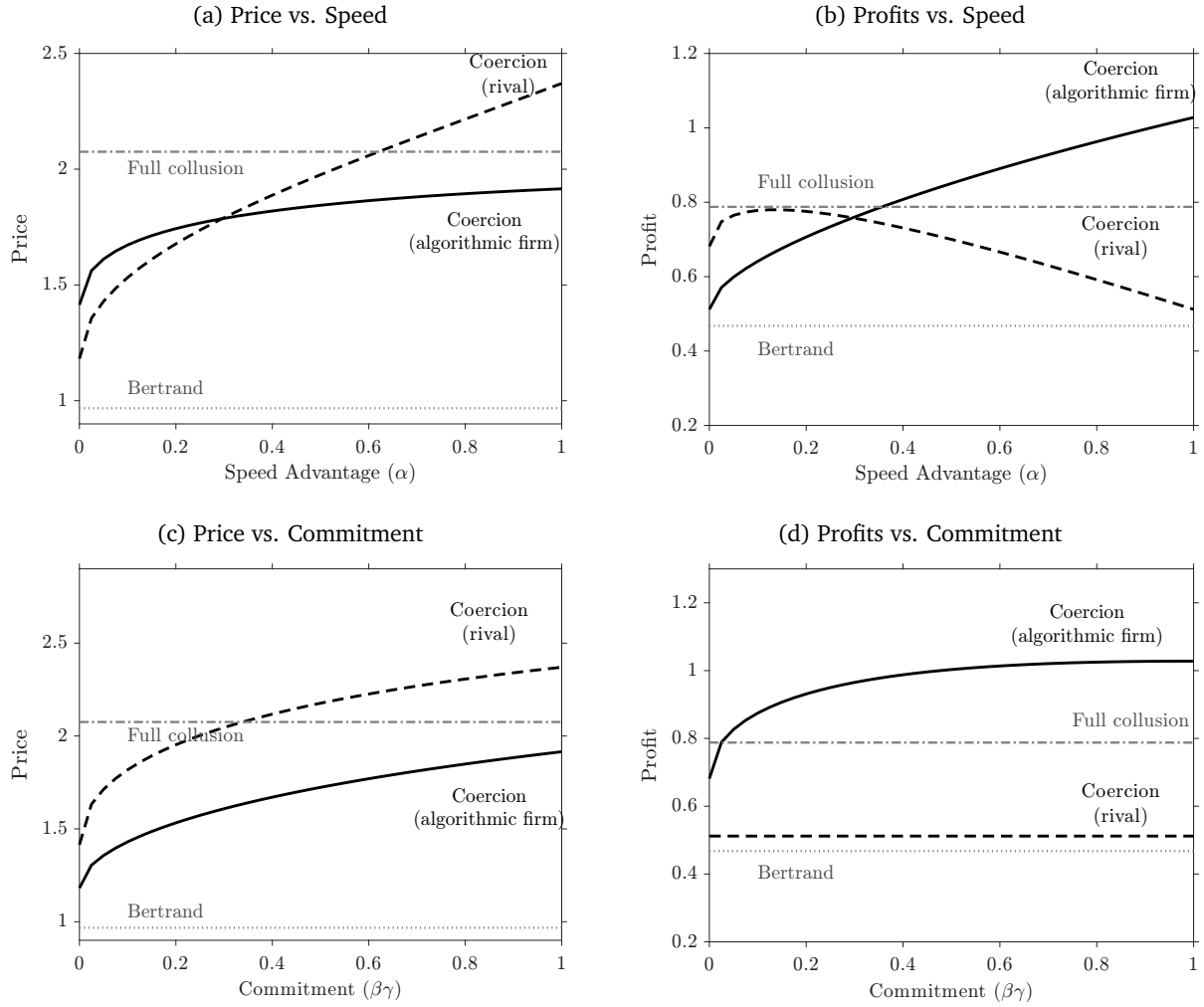
Notes: Panel (a) shows prices and panel (b) shows profits under Bertrand competition with simultaneous pricing, the coercive equilibrium with partial commitment ($\beta\gamma = 0.5$) and speed advantage α, and joint-profit maximization. Demand follows the linear specification in equation (19) with $d = 1$.

Figure A-6: Prices and Profits in Coercive Equilibrium: Intermediate Speed Advantage



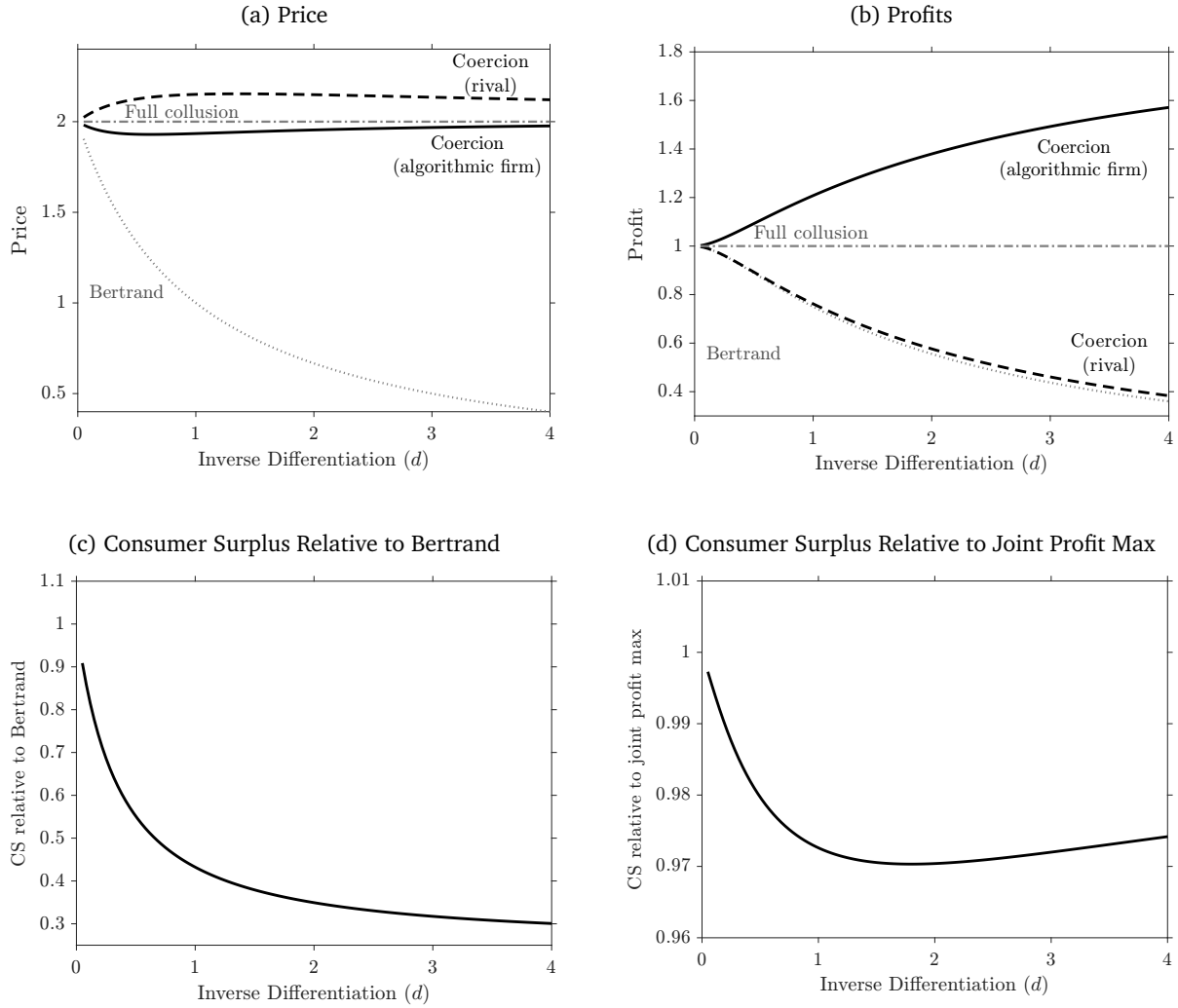
Notes: Panel (a) shows prices and panel (b) shows profits under Bertrand competition with simultaneous pricing, the coercive equilibrium with $\alpha = 0.5$ and commitment $\beta\gamma$, and joint-profit maximization. Demand follows the linear specification in equation (19) with $d = 1$.

Figure A-7: Prices and Profits in Coercive Equilibrium: Logit Demand



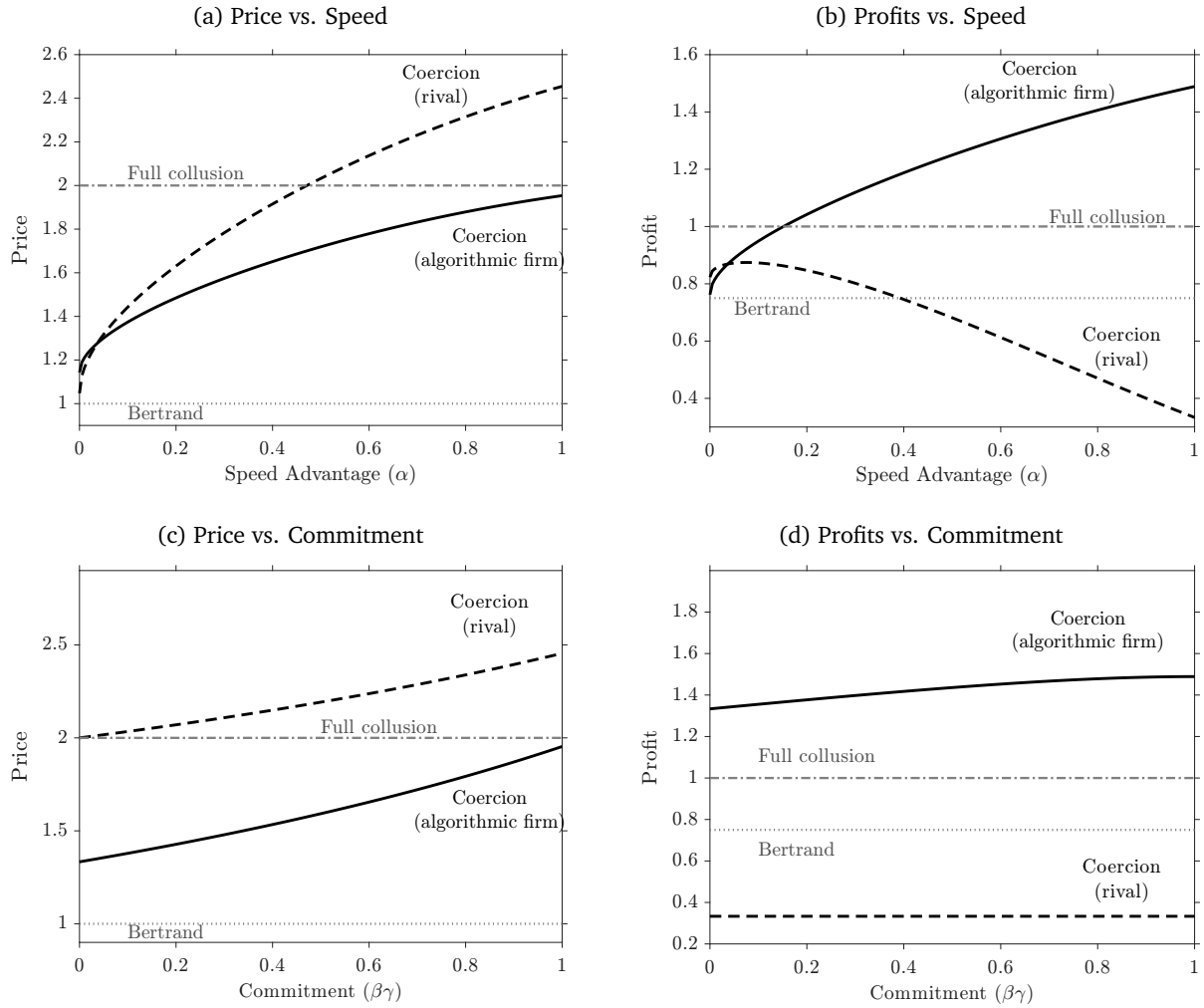
Notes: Panels (a) and (b) show equilibrium prices and profits for different values of the speed advantage α under the coercive equilibrium with full commitment ($\beta\gamma = 1$). Panels (c) and (d) show equilibrium prices and profits under coercion for different values of commitment, $\beta\gamma$,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

Figure A-8: Prices, Profits, and Consumer Surplus in Coercive Equilibrium, by Product Differentiation



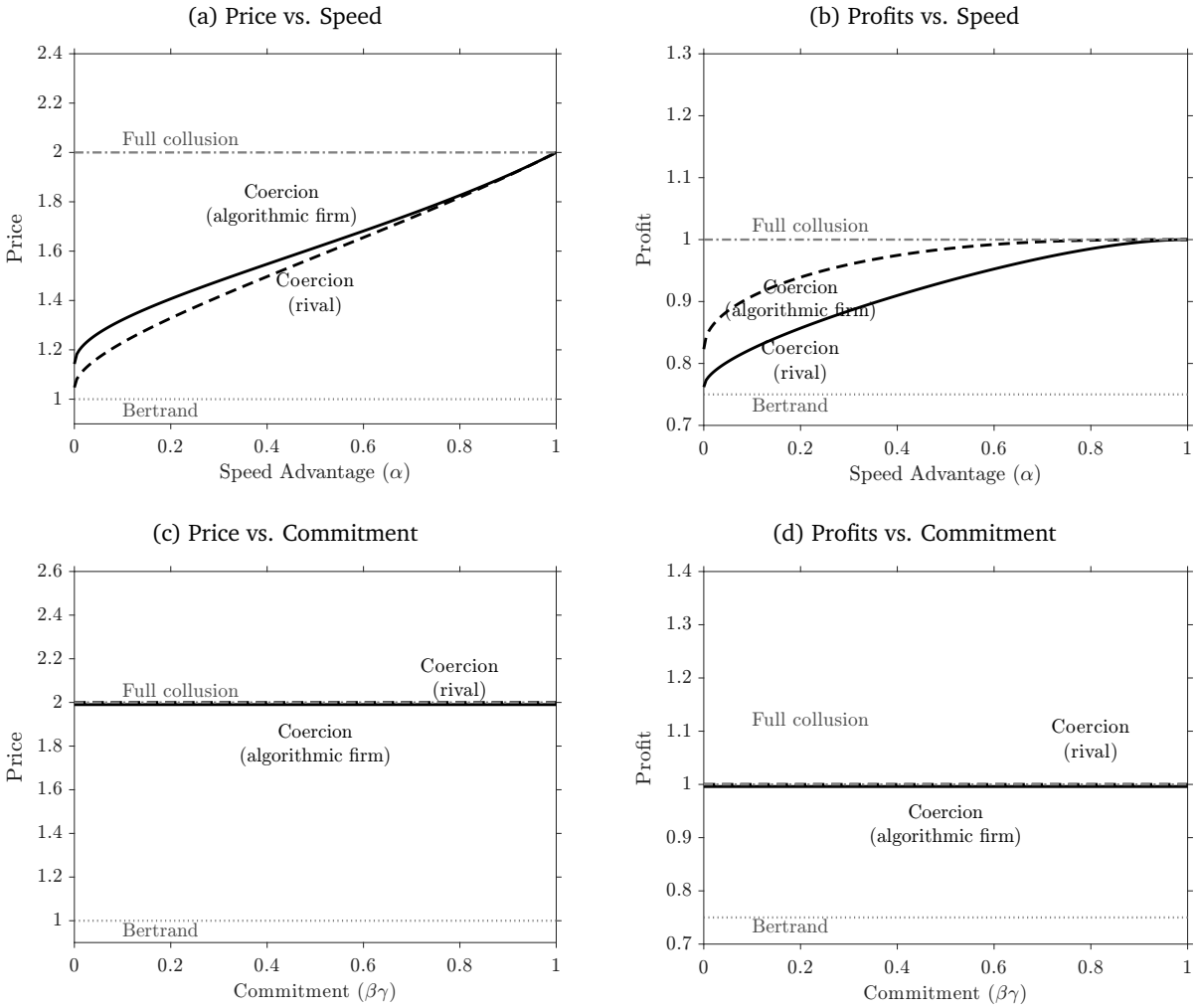
Notes: Panel (a) shows prices and panel (b) shows profits under Bertrand competition with simultaneous pricing, maximal coercion, and joint-profit maximization. Panel (c) shows the ratio of consumer surplus under the coercive equilibrium to consumer surplus under Bertrand competition, and panel (d) shows the analogous ratio relative to joint-profit maximization. A value of one indicates equality with the respective benchmark. The panels show equilibrium outcomes for different values of the inverse product-differentiation parameter d, where larger d corresponds to closer substitutes, using linear demand given by equation (19).

Figure A-9: Prices and Profits in Coercive Equilibrium: Harsh Punishment



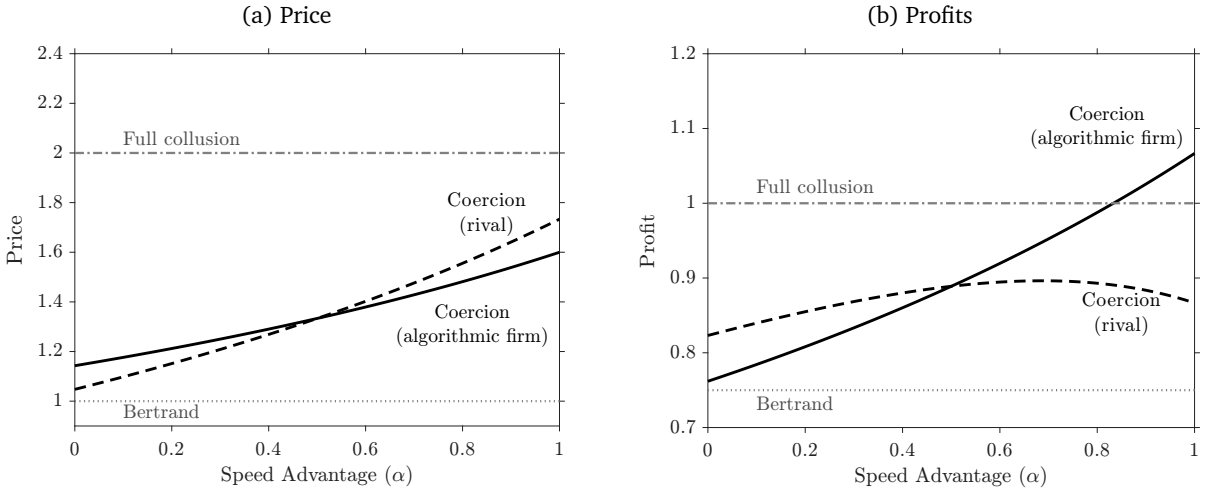
Notes: Panels (a) and (b) show equilibrium prices and profits for different values of the speed advantage α under coercion with full commitment ($\beta\gamma = 1$). Panels (c) and (d) show equilibrium prices and profits under coercion for different values of commitment, $\beta\gamma$, when the faster 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 marginal-cost punishment rather than the static best response. Demand follows the linear specification in equation (19) with $d = 1$.

Figure A-10: Prices and Profits in Coercive Equilibrium: Price Matching Punishment



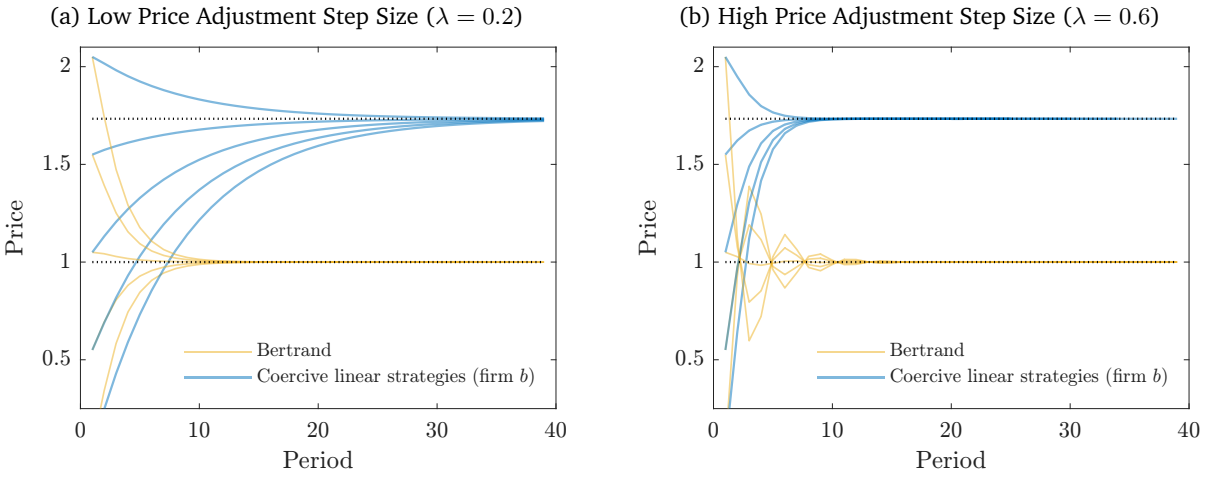
Notes: Panels (a) and (b) show equilibrium prices and profits for different values of speed advantage α under coercion with full commitment ($\beta\gamma = 1$). Panels (c) and (d) show equilibrium prices and profits under coercion for different values of commitment, $\beta\gamma$, when the faster firm 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 the rival's price rather than the static best response. Demand follows the linear specification in equation (19) with $d = 1$.

Figure A-11: 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 for Bertrand competition with simultaneous pricing, the coercive equilibrium in which the faster firm uses a linear punishment rule and has full commitment ($\beta\gamma = 1$), and 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. Demand follows the linear specification in equation (19) with $d = 1$.

Figure A-12: Convergence of Simulated Learning with Linear Pricing Rule versus Simultaneous Pricing



Notes: The figure shows simulated price paths for firm b under gradient learning from five initial price pairs, with $p_b^0 \in \{0, 0.5, 1, 1.5, 2\}$ and $p_b^1 = p_b^0 + 0.05$. For the orange paths, firm b learns while firm a's price is held fixed at the Bertrand–Nash price, $p_a = 1$; these paths converge to $p_b = 1$. The blue paths show the case in which the algorithmic firm has maximum speed advantage ($\alpha = 1$) and uses the linear pricing rule described in Section 5.2. Panel (a) uses a price-adjustment step size of $\lambda = 0.2$, and panel (b) uses $\lambda = 0.6$. Demand follows the linear specification in equation (19) with $d = 1$.

Table A-5: Periods to Convergence to Coercive Outcome

(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: The 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 the first of ten consecutive posted prices within 1% of the equilibrium price for panel (a), within 0.1% for panel (b), and within 0.01% for panel (c). For each value of λ, we simulate price paths from 1,000 initial prices on a grid between 0 and 2 and report the mean, min, max, and percentiles. The simulations assume that the algorithmic firm has maximum pricing speed ($\alpha = 1$) and uses the linear pricing rule described in Section 5.2, and that $d = 1$ under the linear demand system given by equation (19).