Algorithmic Coercion with Faster Pricing[*]

Zach Y. Brown
University of Michigan[†]

Alexander MacKay
University of Virginia[‡]

September 16, 2024

Abstract

We examine a model in which one firm uses a pricing algorithm enabling faster pricing and multi-period commitment. A rival firm sells a differentiated product and simply maximizes static profits. We characterize a *coercive* equilibrium in which the algorithmic firm maximizes its profits subject to incentive compatibility constraints of its rival. By adopting an algorithm that enables (imperfect) commitment and faster pricing, a firm can unilaterally induce substantially higher equilibrium prices. Often, the profits that the firm can obtain are greater than its share of the symmetric collusion profits. The outcomes in the coercive equilibrium can be worse for consumers than collusive outcomes.

Keywords: Pricing Frequency, Pricing Algorithms, Online Competition, Collusion
JEL Classification: L40, D43, L81, L13, L86

[*]We thank Jim Dana, Scott Kominers, Joe Harrington, and Nathan Miller for helpful comments. We are grateful for the research assistance of Harry Kleyer.

[†]University of Michigan, Department of Economics. Email: zachb@umich.edu.
[‡]University of Virginia, Department of Economics. Email: mackay@virginia.edu.

1 Introduction

In classic models of price competition, supracompetitive prices can be sustained when firms are sufficiently forward-looking. For example, in typical models of collusion, all firms understand the dynamic strategies of rivals and value future profits. In the presence of myopic rivals that maximize static profits, collusion cannot be sustained under the usual assumption of simultaneous price setting behavior. Short-sighted behavior by rivals is an important source of fragility in these models.[1]

Recently, however, the use of high speed pricing algorithms has spread to new markets, allowing for the possibility of asymmetries in pricing behavior (Brown and MacKay, 2023). Algorithmic pricing tools often emphasize the importance of reacting faster to competitors' price changes.[3] Furthermore, the fact that pricing algorithms are not continually adjusted implies a level of commitment to a pricing strategy. The adoption of these tools has not been uniform across firms. Recent empirical work has documented that some firms employ high-speed pricing algorithms that have a significant advantage in terms of their ability to observe rivals' prices and the speed in which prices can react.[4]

Motivated by these facts, we consider a model where a single sophisticated firm uses a high-speed algorithm that allows it to update prices more quickly than its rival, and the firm has an (imperfect) commitment to this algorithm across periods. We then ask: what outcomes can this firm achieve unilaterally, even when its rivals are myopic and simply maximize static profits?

In our model, technology is characterized by the algorithm's relative reaction time to its rival's prices and the (probabilistic) rate that it can update its algorithm each period. The naive rival simply maximizes current period profits. Under an assumption that the algorithm punishes deviations by playing the sophisticated firm's one-shot best response function, the equilibrium is unique. The model nests the standard Bertrand and sequential equilibria as special cases. When there is no speed advantage and full commitment, the model yields the sequential equilibrium where the sophisticated firm is the leader. Conversely, when the algorithm can react instantly but there is no commitment across periods, the equilibrium is that where the sophisticated firm is the sequential follower.

In the presence of both a speed advantage and multi-period commitment, the sophisticated firm can obtain prices and profits that are substantially higher than the sequential payoffs. The

[1]To wit, several factors are viewed as important for facilitating collusion, including similar size and costs, predictability of demand, the observability of all rivals prices, and possibility for frequent direct communication.[2]

[3]For instance, a firms offering an algorithmic pricing tool notes that "businesses can compete more effectively by responding quickly" (dealhub.io, 2023). Another advertises that "a fast reaction to your competitors' price variations is essential to be aggressive and competitive in the world of online commerce" (competitoor.com, 2023). One third-party pricing algorithm offers a algorithm 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).

[4]Brown and MacKay (2023) show that the pricing technology for large online retailers varies from once-per-week updates to updates that occur multiple times per hour. The adoption of high-speed pricing algorithms has also been observed in other settings (e.g., Assad et al., 2022; Aparicio et al., 2021).

speed advantage allows the sophisticated firm to coerce its myopic rival into playing higher prices; commitment across periods prevents the sophisticated firm from deviating from its own optimal long-run strategy. Thus, the commitment plays an important role in allowing the sophisticated firm to "tie its hands." The commitment feature also distinguishes our model from earlier work by Brown and MacKay (2023), which did not allow for multi-period commitment.

These results provide some context for an understanding of the potential impact of pricing algorithms. With full commitment, the sophisticated firm can always guarantee at least the sequential (or Stackelberg) payoffs. However, the potential impacts on equilibrium prices can be far greater when the firm also has a speed advantage. Moreover, even with imperfect commitment, faster pricing can still yield equilibrium outcomes where the sophisticated firm earns more than its share of the symmetric collusion profits. In some environments, the fast firm may dictate that the slower firm set prices above the joint profit-maximizing price, leading to a greater deadweight loss and lower welfare than that obtained under joint profit maximization.

Because these outcomes do not depend on cooperative behavior by other firms, they are in some ways more robust than standard models of supracompetitive prices. In the baseline model, the myopic firm is fully informed, but the outcomes can also be generated under alternative assumptions of naive behavior by rival firms. First, the naive firm could be memoryless, in which case it cannot condition its action on past play. Second, the naive firm may not even be aware that it is playing a game against a rival; instead, it may attempt to maximizing a function of its own price only. We also consider the case in which the naive firm is not fully informed, but learns over time.

While we start by examining trigger strategies, the punishment need not be that drastic. Our analysis of pricing rules that are a linear function of the slower rival's price demonstrates that such strategies can lead to collusive prices while potentially raising less antitrust scrutiny and being robust to the use of simple profit optimization by the slower firm.

There is growing concern about collusion in online markets, especially when firms use algorithms (Harrington, 2018). The literature has has largely focused on simultaneous move games, including a literature on collusion with artificial intelligence (Waltman and Kaymak, 2008; Calvano et al., 2020). It is well known that in simultaneous move games, high frequency pricing implies a smaller per period discount rate, making collusion easier to sustain when firms have perfect monitoring (e.g., Abreu et al., 1991). However, asymmetries—typically in terms of costs or demand—are generally believed to make collusion more difficult to sustain (Scherer, 1980; Tirole, 1988). One exception is asymmetries in the discount factor or pricing speed across markets, which can facilitate collusion (Bernheim and Whinston, 1990). 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. Finally, the literature on price leadership and collusion has examined the incentives to collude when one firm announces a price and a rival follows (e.g., Mouraviev and Rey, 2011). In previous work, Brown and MacKay (2023) examine competitive (Markov perfect) outcomes when firms may differ in the speed at which they set prices. We are not aware of other research that examines collusion in settings where firms differ in their pricing frequency.[5]

The paper proceeds as follows. We introduce the model in Section 2. In Section 3, we discuss the equilibrium concept and provide benchmark cases. We characterize the general problem and provide examples in Section 4. In Section 5, we introduce the idea of learning for the naive firm, and we show how the sophisticated firm can obtain the collusive or coercive equilibria even with simple linear strategies. Section 6 indicates how our results can be generalized to N-firm oligopolies where a single firm has a pricing speed advantage and the other $N-1$ firms are naive.[6] Section 7 concludes.

2 Model

We analyze a duopoly model in which a sophisticated firm can commit to an algorithm that automatically updates prices, while its naive rival is myopic and has simple price-setting technology. We discuss how our results may be generalized to an N-firm oligopoly in Appendix 6.

2.1 Environment

Two firms, a and b, compete in an environment where profits are determined by prices, (p_a, p_b). Firm a is sophisticated and has a pricing algorithm, while firm b is naive and does not have an algorithm. Time is continuous and is given by $t \in [0, \infty]$, while periods are defined by discrete intervals indexed by $t \in \{0, 1, 2, ...\infty\}$. We define the length of a period as the frequency that firms without algorithms can update prices. Thus, firm b can update prices at the beginning of each period, $t \in \mathbb{N}_0$.

Demand arrives in continuous time. The instantaneous profit flow function for firm i is time-invariant and is given by $\pi_i(p_i, p_{-i})$. We assume the profit functions are quasi-concave and have a unique maximum with respect to a firm's own price.

[5]Our paper also relates to a previous literature examining games in which a single long-run player faces a succession of repeated short-run players. The classic application is one in which an incumbent faces a new (short-run) potential entrant in each period, as studied by Milgrom and Roberts (1982) and Kreps and Wilson (1982). Fudenberg and Levine (1989) and Fudenberg et al. (1990) provide folk-theorem style analysis for feasible payoffs for a general class of games with a single long-run player. Our work contributes to this literature by considering repeated games of price competition and the advantage conferred to a single firm through the adoption of a pricing algorithm. In our setting, we consider repeated interactions with the same rival and interpret short-run behavior as arising from naive play.

[6]When the faster firm can observe each of the prices of its rivals, it can specify a punishment to each rival's possible deviation that ensures no rival deviates in equilibrium.

2.2 The Sophisticated Firm

The sophisticated firm a maximizes discounted future profits over an infinite horizon. Firm a has a discrete discount rate of profits in future periods given by $\beta > 0$.[7]

At time $t = 0$, and at future opportunities indexed by τ, firm a chooses a pricing algorithm, $\mathcal{A}_{\tau}^{\alpha,\gamma}(p_{bt}, x_t, t)$. The algorithm is characterized by speed and commitment parameters, α and γ, which we describe below. The algorithm sets prices according to

$$p_{at} = \mathcal{A}_{\tau}^{\alpha,\gamma}(p_{bt}, x_t, t) \equiv \begin{cases} \rho_\tau(x_t) & t - \lfloor t \rfloor \leq \alpha \\ \sigma_\tau(p_{bt}, x_t) & t - \lfloor t \rfloor > \alpha \end{cases} \tag{1}$$

where $\lfloor t \rfloor$ is the floor function that yields the greatest integer less than t, i.e., the beginning of the period. The algorithm is characterized by an initial price-setting function $\rho_\tau(x_t)$ and an update function $\sigma_\tau(\hat{p}_{bt}, x_t)$ that can depend on observable state variables, x_t. The initial-price setting component $\rho_\tau(x_t)$ is used to set prices at the beginning of each period, e.g., $p_{a0} = \rho_0(x_0)$. Within each period, after the elapsed interval $\alpha \in [0, 1]$, the fast firm can observe p_{bt} and adjust its price to according to the update function, e.g., $p_{a\alpha} = \sigma_0(\hat{p}_{bt}, x_0)$. We call α the *reaction time*, and it captures the algorithm's ability to update prices more frequently.

These features reflect features of software that is used to update prices automatically. At the beginning of each period, the algorithm may update the price in response to new information captured by x_t. During these updates, firm b can also change its price, but the decision of firm b is not known in advance by the algorithm. However, the algorithm can observe p_{bt} chosen in that period and then update with a lag α. One can interpret our equilibrium analysis as conditional on a sequence of state variables, therefore, going forward, we suppress x_t in our notation.

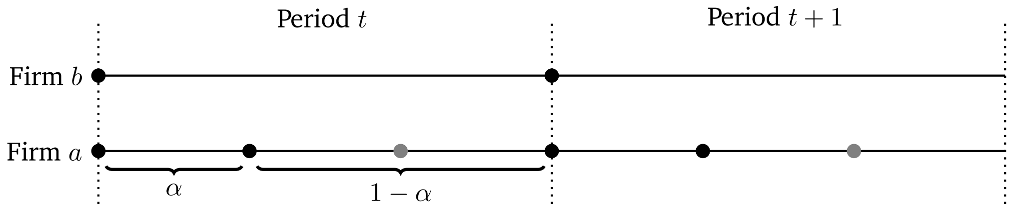
Figure 1 shows the timing of price adjustments. Both firms can adjust prices at t, firm a can adjust at $t + \alpha$, and firm b must wait fraction $1 - \alpha$ until $t + 1$ to adjust its price in response. When $\gamma = 0$, the model nests a standard simultaneous move game when $\alpha = 1$ and a sequential price-setting game where firm b moves first when $\alpha = 0$.[8] In practice, differences in pricing frequency may provide the faster firm multiple opportunities to adjust prices in each period. However, given the assumptions about our environment, only the first opportunity to adjust price within the period is consequential. In this way, α can be interpreted as the lag between firm b setting price and the first opportunity for firm a to readjust its price in response.[9]

[7] For simplicity, we will assume that 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, this is without loss of generality since including within-period discounting is equivalent to defining alpha as the subjective reaction time. Consider any α and any instantaneous within-period discount rate ν. Then there exists an objective (non-discounted) reaction time $\tilde{\alpha}$ such that $\frac{\int_0^{\tilde{\alpha}} e^{-\nu t} dt}{\int_0^1 e^{-\nu t} dt} = \frac{1 - e^{-\nu \tilde{\alpha}}}{1 - e^{-\nu}}$. For a given δ, the mapping of α to $\tilde{\alpha}$ is one-to-one.

[8] We discuss these connections in more detail in Section 3.

[9] In the case of high-speed pricing algorithms, α is determined by the time required for software to observe a rival's price and update price in response. In online retail markets, a large pricing speed advantage has been observed for

Figure 1: Timing with Differences in Pricing Frequency



Notes: Figure shows timing of pricing when firm a, the sophisticated firm, can react with lag α. Circle markers represent opportunities to adjust prices. The gray circle represents an opportunity to adjust price that is inconsequential in the model.

The firm has an indefinite commitment to the algorithm over future periods, which is captured by γ. When the firm is not updating the algorithm, the price is set automatically by the algorithm at each price change opportunity, e.g., $t \in \{\alpha, 1, 1+\alpha, 2, 2+\alpha ...\}$. At the beginning of each period, with probability γ, firm remains committed to the algorithm and cannot manually update the algorithm or its price. With probability $1-\gamma$, the firm can update the algorithm, at which point it also picks a new initialization price.

The sophisticated firm's decisions can be expressed as a dynamic problem:

$$V_0(t) = \max_{\mathcal{A}|p_{bt}} \alpha\pi_a(\rho, p_{bt}) + (1-\alpha)\pi_a(\sigma(p_{bt}), p_{bt}) + \beta\gamma V_1(t+1, \mathcal{A}) + \beta(1-\gamma)V_0(t+1) \quad (2)$$

$$V_1(t, \mathcal{A}) = \alpha\pi_a(\rho, p^*_{bt}) + (1-\alpha)\pi_a(\sigma(p^*_{bt}), p^*_{bt}) + \beta\gamma V_1(t+1, \mathcal{A}) + \beta(1-\gamma)V_0(t+1) \quad (3)$$

starting with $t = 0$ and for each integer t thereafter. $V_0(t)$ provides the value function when firm a can update the algorithm, and $V_1(t, \mathcal{A})$ provides the value function when the firm is committed to algorithm $\mathcal{A}$. Here, p^*_{bt} gives the optimal reaction by firm b to $\mathcal{A}$. Firm a can anticipate the optimal response of firm b in future periods while it remains committed to the algorithm.

2.3 The Naive Rival

We now discuss the naive rival. In the model, we assume firm b is myopic and sets prices to maximize current period profits only.

Assumption A1. *In each period $t \in \{0, 1, 2, ...\}$, firm b solves the problem*

$$\max_{p_b|\mathcal{A}_\tau} \alpha\pi_b(p_b, \rho_\tau) + (1-\alpha)\pi_b(p_b, \sigma_\tau(p_b)) \quad (4)$$

certain firms (implying α is small). Brown and MacKay (2023) consider the differences in pricing frequency among major U.S. online retailers. A slower retailer updates prices once at the beginning of each week, whereas another retailer updates the price once each day. At the beginning of the week, the firms set prices simultaneously. Given that the fast firm can respond to the slow firm's price the following day, $\alpha = 1/7$ in this case.

We initially assume that the firm understands $\mathcal{A}_\tau$ and internalizes the sophisticated firm's reaction within the period. We call firm b naive in that it does not anticipates what happens across periods. We make this restriction so that the rival acts in its own short-run best interest. Such "short-termism" may arise due to, e.g., misalignment between managerial incentives and shareholders. As we will be made clear in the sections that follow, this behavior by a naive rival will make standard collusive equilibria unattainable. The naive firm treats each period as a one-shot game and acts as if it is maximizing its static period profits.

Two other behavioral assumptions can also generate this naive behavior. First, the naive firm may be memoryless, in that is not able to condition its actions on the history of past play. Such an assumption would eliminate the ability of the firm to form contigent strategies that punish the sophisticated firm. Second, the slow firm may be non-strategic, in that it is unaware that it is playing a dynamic game with rival firms. For example, it may (falsely) believe that it is maximizing the profit function $\bar{\pi}_b(p_b)$ that does not depend on the strategic variable p_a. In these two cases, the slower firm's solution to the repeated game is equivalent to the myopic solution, even though the slow firm values future profits. Thus, all three assumptions give rise to behavior that can be captured by the myopic case.

Myopia and memoryless behavior are distinguished from the non-strategic behavior in that the slow firm may be fully aware of the actions of the fast rival and be able to predict its actions. In these cases, the slow firm is limited in terms of how far forward or backward it looks in time when deciding its actions.

In the non-strategic case, a firm does not internalize the fact that it is playing a game with strategic rivals. This is known to be the case in various oligopolistic markets, such as as airline ticket pricing.[10] The slow firm may be modeled as simply trying to maximize profits as a function of a single input (its own price). Consider, for example, a firm that sets prices at the beginning of the week, observes profits at the end of the week, and attempts to learn which price yields the highest profits. The firm specifies its profit function as $\pi_b(p_b)$. In truth, the profit function depends also on the price of the faster firm, i.e., $\pi_b(p_b, p_a)$, but the slow firm is not aware of this. The faster firm can vary p_a throughout the week to manipulate the slow firm into learning that the optimal p_b is a particular target price, for example, the fully collusive price.

Standard models of collusion are not attainable in equilibrium if any of the above behaviors are present. All firms (1) must have positive valuations of future periods, (2) must be able to condition on the history of play, and (3) must be aware of the strategies of rival firms. We will show this by examining benchmark equilbria in the following section.

In Section 5, we relax the informational assumptions and consider the case in which the slow firm learns about the algorithm over time.

[10]For instance, Hortaçsu et al. (2022) document that a major airline does not internalize the existence of any competition when setting ticket prices regardless of the market structure.

3 Equilibrium Concept and Benchmark Cases

3.1 Equilibrium Concept

Here, we define the equilibrium concept, and we explore benchmark equilibria under different assumptions about frequency and commitment. These benchmark cases help to build intuition for these features and motivate the general analysis that we present in Section 4.

Equilibrium is characterized by a sequence of realized algorithms $\{\mathcal{A}_t\}$ and prices $\{p_{bt}\}$ such that equations (2), (3), and (4) are satisfied for all t. We use a standard Nash equilibrium condition. When choosing an algorithm, firm a takes as given the current price of firm b. When choosing a price, firm b takes as given the algorithm chosen by a.

With time-invariant profit functions, he problem is stationary. We can exploit the fact that $V_0(t) = V_0(t')$ for $t, t' \in \mathbb{N}$ to express the sophisticated firm's problem from equations (2) and (3) as:

$$\begin{aligned}\tilde{V}(t) = &\max_{\mathcal{A}|p_{bt}} \alpha\pi_a(\rho, p_{bt}) + (1-\alpha)\pi_a(\sigma(p_{bt}), p_{bt}) \\ &+ \sum_{s=t+1}^{\infty} (\beta\gamma)^s \left(\alpha\pi_a(\rho, p_{bt}^*) + (1-\alpha)\pi_a(\sigma(p_{bt}^*), p_{bt}^*)\right)\end{aligned} \tag{5}$$

where $\tilde{V}(t) = \frac{1-\beta}{1-\beta\gamma} V_0(t)$. Going forward, we will make use of the fact that $\sum_{s=t+1}^{\infty} (\beta\gamma)^s = \frac{\beta\gamma}{1-\beta\gamma}$ to simplify notation.

Because the profit function π_a is quasi-concave (with a unique maximum), $\tilde{V}(t)$ is quasi-concave. With mild restrictions on $\sigma(\cdot)$, the equilibrium is unique and coincides with Markov Perfect equilibrium.[11] The naive rival does not account for future profits or respond to the history of play. Thus, its presence eliminates a large class of equilibria that can be supported in repeated games.

To see this, consider the case when $\alpha = 1$ and $\gamma = 0$, so that the algorithm provides no commitment and no speed advantage. The objective functions become

$$\text{Firm } a : \max_{\mathcal{A}|p_b t} \pi_a(\rho, p_{bt}) \tag{6}$$

$$\text{Firm } b : \max_{p_b|\mathcal{A}_\tau} \pi_b(p_b, \rho_\tau) \tag{7}$$

which corresponds to the one-shot simultaneous price-setting game. Thus, the only subgame perfect equilibrium is the Bertrand-Nash equilibrium. Though dynamic price setting games may, in general, yield multiple equilibria, the presence of the myopic firm greatly reduces the set of outcomes that can be sustained in equilibrium.

[11] Specifically, there is a unique $\sigma^*(\cdot)$ that is a (weakly) dominant strategy for firm a, even when other update functions also are consistent with equilibrium behavior, including those that yield the same outcomes.

We now consider two benchmark cases. First, we consider the case where the algorithm has some commitment but no speed advantage. Next, we consider the case with a speed advantage but no multi-period commitment. This second case corresponds to the asymmetric commitment model analyzed in Brown and MacKay (2023).

3.2 Multi-Period Commitment Only

In some settings, the algorithm may provide a commitment advantage but no speed advantage. This is the case in the literature on learning algorithms and competition (e.g., Calvano et al., 2020; Johnson et al., 2021; Asker et al., 2022), which have firms that set prices simultaneously. The analysis here can be roughly thought of as an extension of these models, where one firm can endogenously choose the optimal algorithm and the other firm uses learning algorithm that recovers its the true payoffs. We will not discuss learning here but instead describe the long-run payoffs.

With no speed advantage, $\alpha = 1$. Firm a has the objective function

$$\max_{\mathcal{A}|p_{bt}} \underbrace{\pi_a(\rho, p_{bt})}_{\substack{\text{Simultaneous}\\ \text{Pricing Incentive}}} + \underbrace{\frac{\beta\gamma}{1-\beta\gamma}\pi_a(\rho, p^*_{bt})}_{\substack{\text{Leader}\\ \text{Pricing Incentive}}} \tag{8}$$

while firm b maximizes $\max_{p_b|\mathcal{A}_\tau} \pi_b(p_b, \rho_\tau)$. The objective function for firm a differs from that of the one-shot benchmark due to the term $\frac{\beta\gamma}{1-\beta\gamma}\pi_a(\rho, p^*_{bt})$, which is positive as long as $\beta\gamma > 0$. The sophisticated firm balances the profits in the current period, conditional on p_{bt}, against the profits in future periods where the naive rival might update its price.

The second component in the objective captures the pricing incentive of a leader in a sequential pricing game. In equilibrium, p^*_{bt} will be given by firm b's static best response function, $p^*_{bt} = R_b(\rho)$. Because a anticipates profits in future periods that it is commuted to the algorithm, it internalizes the reaction of its rival. In the limiting case where firm a is infinitely patient ($\beta = 1$) and has perfect commitment ($\gamma = 1$), the outcome is equivalent to a sequential price-setting game in which firm a is the (Stackelberg) leader. For example, in a setting in which the naive rival sets prices frequently (say once per day), and firm a updates its algorithm once per month on average, the outcome may be similar to sequential pricing, as firm a will put a lot of weight on the evolution of profits over future periods.

More generally, the equilibrium can be characterized as follows:

Proposition 1. *When the algorithm enables commitment ($\gamma > 0$) but no speed advantage ($\alpha = 1$), the equilibrium lies along the naive firm's best response function between the simultaneous price-setting equilibrium and the sequential price setting equilibrium where the sophisticated firm is the leader.*

Proof. This follows directly from the fact that the sophisticated firm maximizes a weighted sum of the simultaneous and sequential profits. □

When firms produce substitute goods and prices are strategic complements, both firms realize higher prices compared to the price-setting (Bertrand-Nash) equilibrium. This follows similar logic of Proposition 2 of Brown and MacKay (2023). In contrast to the asymmetric commitment model of Brown and MacKay (2023), when firms a and b have identical profit functions, the firm with the algorithm has a higher price than its naive rival.

3.3 Faster Pricing Only

We now consider the case in which firm a has no multi-period commitment ($\gamma = 0$) but the algorithm enables faster pricing updates ($\alpha < 1$). The period t objective function for firm a when it can update its algorithm becomes

$$\max_{\mathcal{A}|p_{bt}} \underbrace{\alpha\pi_a(\rho, p_{bt})}_{\substack{\text{Simultaneous} \\ \text{Pricing Incentive}}} + \underbrace{(1-\alpha)\pi_a(\sigma(p_{bt}), p_{bt})}_{\substack{\text{Follower} \\ \text{Pricing Incentive}}} \tag{9}$$

while the objective function for firm b remains as given in equation (4).

This problem is equivalent to the asymmetric commitment model analyzed by Brown and MacKay (2023). Following that analysis, it is weakly dominant for firm a to choose an update function that corresponds to its static best-response function, $\sigma(\cdot) = R_a(\cdot)$. In Section 4, we make a specific restriction on $\sigma(\cdot)$ that nest this as a special case and yields a unique equilibrium for the more general problem.

In contrast to the above case with only multi-period commitment, the sophisticated firm balances the simultaneous price-setting incentive with the sequential price-setting incentive where it acts as the *follower*. Thus, following Proposition 2 from Brown and MacKay (2023), the equilibrium lies on the sophisticated firm's best response function, between the simultaneous and the sequential equilibrium. The parameter α indicates how much weight the naive firm puts on the period before the algorithm update. As above, this results in higher prices for both firms when the products are substitutes and prices are strategic complements. However, this case will yield lower prices for the sophisticated firm, instead of higher prices, when the firms have identical profit functions.

In the limiting case where $\alpha = 0$, the algorithm yields the sequential price-setting equilibrium, with firm b acting as the leader and firm a acting as the follower. Thus, the two features of the algorithms we study—speed and multi-period commitment—can generate sequential equilibria where the sophisticated firm takes on either the leader or follower role.

Figure 2: Benchmark Equilibria, with Examples

		Commitment	
		$\beta\gamma = 0$ (Low)	$\beta\gamma = 1$ (High)
Reaction Time	$\alpha = 1$ (Slow)	Simultaneous Bertrand-Nash $(p_a, p_b) = (1.00, 1.00)$	Sequential, Firm a is Leader $(p_a, p_b) = (1.14, 1.05)$
	$\alpha = 0$ (Fast)	Sequential, Firm a is Follower $(p_a, p_b) = (1.05, 1.14)$	Maximal Coercion $(p_a, p_b) = (1.93, 2.15)$

3.4 Discussion and Examples

We have thus far considered three benchmark cases. First, we demonstrated that with simultaneous price-setting and no multi-period commitment, the unique equilibrium is the one-shot Bertrand-Nash equilibrium, corresponding to the Markov Perfect equilibrium with simultaneous pricing. Then, we showed that full commitment or instantaneous reactions can yield sequential price-setting equilibria, though the roles of the algorithmic firm depend on the features of the algorithm.

In Section 4, we consider the general case where $\alpha < 1$ and $\beta\gamma > 0$. We call these equilibria *coercive*. In general, these need not lie between any of the above benchmark cases, but instead expand the set of possible feasible outcomes. An extreme case is one in which $\alpha = 0$ and $\beta\gamma = 1$, in which firm a is infinitely patient, has full commitment, and can instantaneously react to the price of firm b in any period. We call this limiting case *maximal coercion*.

To illustrate these cases, here and elsewhere in the paper, we use a simple symmetric linear demand system given by

$$D_i(p_i, p_{-i}) = 1 - \left(\frac{1}{4} + \frac{d}{2}\right) p_i + \frac{d}{2} p_{-i} \tag{10}$$

where $d \geq 0$ is an inverse measure of product differentiation. This demand system can be derived from the quasilinear quadratic utility model (Singh and Vives, 1984). The goods do not compete when $d = 0$ and are perfect substitutes when $d = \infty$. Without loss of generality, marginal costs are normalized to zero.

Figure 2 summarizes the benchmark cases for the limiting values of α and $\beta\gamma$. Equilibrium prices are reported for the demand system with $d = 1$. The implications for profits are substantial. The sophisticated firm earns profits of 0.75 in the Bertrand-Nash benchmark, 0.76 as the sequential leader, 0.82 as the sequential follower, and 1.21 in the coercive equilibrium. Thus, moving from the Betrand-Nash benchmark to maximal coercion increases the profits for firm a

by 61 percent. By comparison, the collusive price levels are given by $(p_a, p_b) = (2, 2)$, yielding profits of 1.00 to each firm. With maximal coercion, the sophisticated firm can obtain higher profits for itself than the collusive outcome.

4 Algorithmic Coercion

4.1 General Problem

We now consider the general case of algorithmic coercion. We reformulate equation (5) as a constrained optimization problem. We ask whether a target price pair (p_a^T, p_b^T) can be maintained in equilibrium.

Specifically, we have firm a choose the target prices (p_a^T, p_b^T) that maximizes its discounted profits, subject to the algorithm technology and the incentive-compatibility constraints for both firms. This yields the following objective:

$$\max_{(p_a^T, p_b^T)|p_{bt}} \alpha\pi_a(\rho, p_{bt}) + (1-\alpha)\pi_a(\sigma(p_{bt}), p_{bt}) + \frac{\beta\gamma}{1-\beta\gamma}\pi_a(p_a^T, p_b^T) \tag{11}$$

$$\text{s.t.} \quad (i) \quad \rho = p_a^T \tag{12}$$

$$(ii) \quad \sigma(p_b) = \begin{cases} p_a^T & \text{if } p_{bt} = p_b^T \\ P_a(p_b) & \text{if } p_{bt} \neq p_b^T \end{cases} \tag{13}$$

$$(iii) \quad p_b^T = \operatorname*{argmax}_{p_b|\mathcal{A}} \alpha\pi_b(p_b, \rho) + (1-\alpha)\pi_b(p_b, \sigma(p_b)) \tag{14}$$

which is obtained by plugging in the target prices into equation (5).

The objective function is specified in terms of three constraints: (i) firm a chooses the initial price ρ to be equal to its target price, (ii) the update function follows provides firm a's target price as long as firm b follows its target price, and (iii) the target price for firm b satisfies its incentive-compatibility constraint.

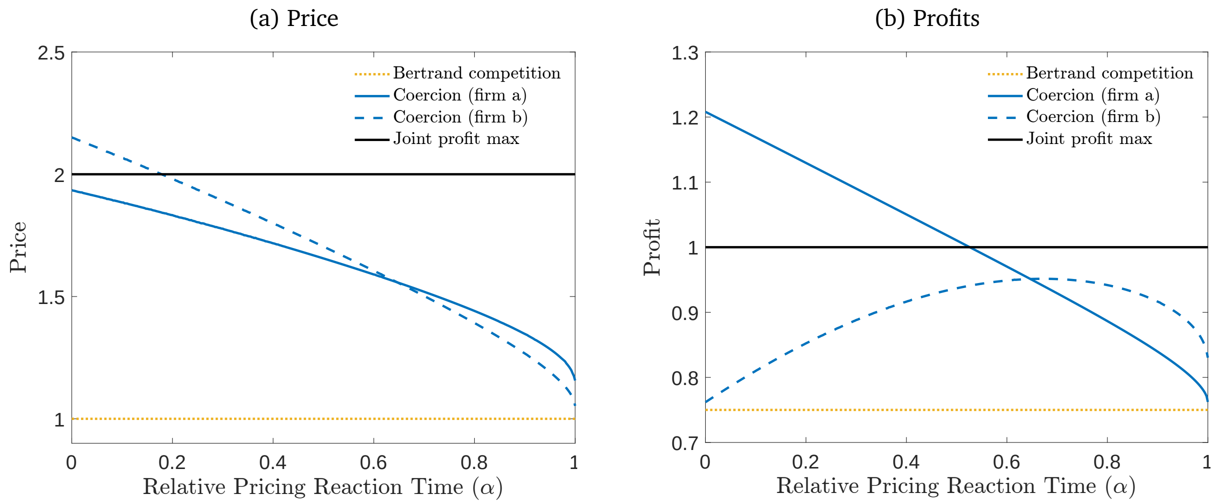
When firm b does not choose $p_{bt} = p_b^T$, the update function of the algorithm follows a potentially arbitrary punishment function, $P_a(p_b)$. Here, we assume that the punishment function is simply the best-response function for firm a:

Assumption A2. *The punishment function is equal to firm a's static best-response function, $P_a(\cdot) = R_a(\cdot)$.*

Though one could consider more extreme forms of punishment, it is typical in the literature on collusion to assume punishment strategies that are consistent with short-run, non-cooperative behavior. We follow that convention here.

There is a unique Markov Perfect equilibrium characterized by the choice of $\mathcal{A}$ and p_b that satisfy the above conditions. Uniqueness is obtained under assumptions A1 and A2 when the

Figure 3: Prices and Profits with Coercive Trigger Strategies and Full Commitment



Notes: Panel (a) shows prices and panel (b) shows profits under Bertrand competition with simultaneous pricing, coercion in which the faster firm uses a trigger strategy to maximizes its own profits with full commitment ($\beta\gamma = 1$), and joint profit maximization. The firms simultaneously set prices at the beginning of the period and then the fast firm can adjust price with reaction time α. Assumes $d = 1$ under linear demand given by equation (10).

profit functions are well-behaved.[12]

4.2 Equilibrium Outcomes

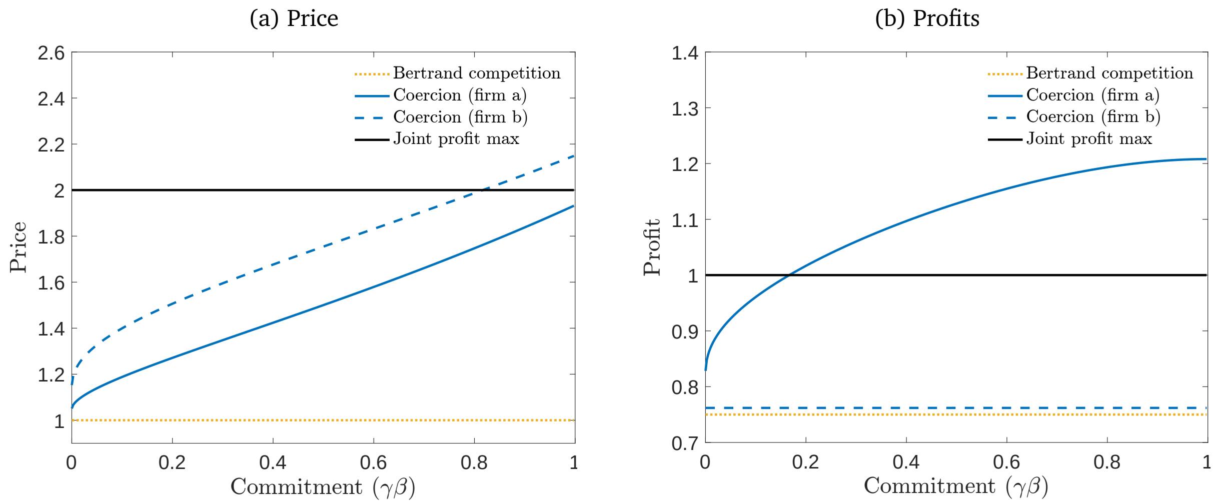
Figure 3 illustrates equilibrium prices (panel (a)) and profits (panel (b)) in the case with full commitment ($\beta\gamma = 1$) and different values for α. The equilibria are generated from the linear demand system with a differentiation parameter of $d = 1$. The solid blue line represents the algorithmic firm, while the dashed blue line represents the naive firm. For comparison, we also plot the joint profit maximizing prices and profits (black line) and the prices and profits for Bertrand competition (yellow dotted line).

Depending on the relative pricing reaction time, there are three regions that determine the relative prices of the slow and fast firm. With very fast pricing, the fast firm can coerce its slower rival into setting prices above the fully collusive price. The fast firm can then undercut this price and earn higher profits. For the intermediate range of pricing speed, the fast firm is able to incentivize the slower rival to set a higher price than the fast firm but cannot coerce the slower rival to set a price above the fully collusive price.

When the fast firm does not have much of speed advantage, it is instead optimal for the fast firm to target a higher price than the slow firm and to use its threat of punishment to prevent the slow firm from lowering its price further. With no speed advantage, we obtain the

[12]Under more general forms for the punishment function, multiple equilibrium can be obtained; thus, A2 can alternatively be viewed as a device for equilibrium selection.

Figure 4: Prices and Profits with Coercive Trigger Strategies by Level of Commitment



Notes: Panel (a) shows prices and panel (b) shows profits under Bertrand competition with simultaneous pricing, coercion in which the faster firm uses a trigger strategy to maximizes its own profits with maximum speed ($\alpha = 0$), and joint profit maximization. The firms simultaneously set prices at the beginning of the period and then the fast firm can adjust price with reaction time $\alpha = 0$. Assumes $d = 1$ under linear demand given by equation (10).

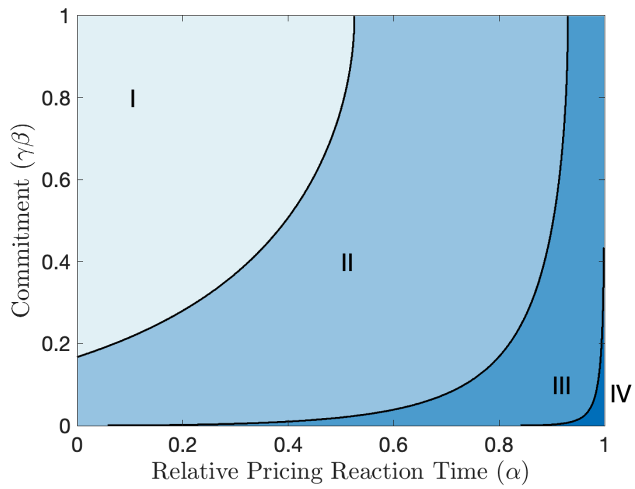
sequential equilbrium as discussed in Section 2. In all cases, prices of both firms are above the Bertrand price.

When the fast firm's pricing speed advantage is small, the slower firm earns greater profits than the fast firm. Around $\alpha = 0.62$, the lines for firm a and firm b intersect. These values reflect the case when the incentives from commitment and the pricing advantage balance each other out, yielding a symmetric increase in prices and profits. Such cases coincide with a symmetric partial collusion equilibrium.

Figure 4 shows the equilibrium cases when commitment ($\beta\gamma$) varies and $\alpha = 0$. For these cases, firm a always prices lower than firm b and earns greater profits. Panel (b) shows that firm b only earns profits approximately equal to the competitive profits in all cases. Thus, firm a can use the threat of an immediate reaction to incentivize firm b to set higher prices and extract all of the resulting producer surplus. The extent to which firm a can do this in Markov Perfect equilibrium depends on its ability to commit. When commitment is low, firm a has a short-run incentive to reduce prices given the high prices of its rival.

There is an important interaction between commitment and reaction time. With only commitment or only fast reactions, firm a is bounded in its ability to raise prices to the sequential prices and payoffs. Appendix Figures A-3 and A-4 illustrate the cases of $\beta\gamma = 0$ and $\alpha = 0$, respectively, showing a modest increase in prices and profits. In the presence both features—as illustrated by moving right to left in Figure 3 or left to right in Figure 4 —a sophisticated firm can obtain substantially higher prices and profits. These strong interaction effects persist at

Figure 5: Firm a Profits by Pricing Speed and Commitment



Notes: Shows profit regions of firm a for different values of reaction time (α) and commitment ($\beta\gamma$). Region I indicates profits greater than under joint profit maximization (splitting collusive profits), region II indicates profits greater than sequential follower, region III indicates profits greater than sequential leader, and region IV indicates profits greater than under Bertrand competition. Assumes $d = 1$ under linear demand given by equation (10).

intermediate values of commitment and reaction time, as illustrated in Appendix Figures A-1 and A-2.

Figure 5 plots the profits obtained by firm a under all combinations of commitment and reaction time. Region I indicates profits greater than symmetric collusion, region II indicates profits greater than those obtained by a sequential follower, region III indicates profits greater than those obtained by a sequential leader, and region IV indicates profits greater than under Bertrand competition. Appendix Figure A-5 shows the same for firm b.

For this demand system, firm a can obtain profits greater than the sequential leader for almost all values of α and $\beta\gamma$. Moreover, many combinations of commitment and speed can allow the sophisticated firm to obtain greater profits than the symmetric collusion payoffs.

The equilibrium yields lower consumer surplus than under Bertrand competition. For most cases with this demand system, consumer surplus is lower than under the sequential equilibrium. When $\beta\gamma$ is close to 1 and α is close to zero, consumer surplus is lower than that obtained under symmetric collusion (joint profit maximization). We provide a plot of consumer surplus in Appendix Figure A-6

This analysis highlights how two features of algorithms—speed and commitment—can yield higher prices and provide a substantial advantage to the adopting firm. This is the case even in the somewhat extreme conditions we assume in our baseline model, where the rival is only acting it its short-run best interest.

5 Incorporating Learning

In the above analysis, we show how a faster firm may unilaterally implements supracompetitive prices when the slow firm understands (explicitly or implicitly) the potential punishment strategy and resulting profits. We now consider the case in which the slow firm is potentially uninformed about the strategy used by the fast firm or its own profit function. Instead, the firm learns over time using an optimization algorithm. We ask whether the fast firm can induce higher prices without announcing a strategy.

5.1 Simple Learning and Linear Strategies

In many settings, it may be reasonable to assume that an equilibrium arises not from introspection by the players but rather from an iterative process of adaptive learning (Fudenberg and Levine, 2016). Models of learning in economic games include fictitious play, reinforcement learning, and gradient learning. In a game in which two firms play a simultaneous pricing game and use these learning strategies to maximize static profits, the Betrand-Nash equilibrium is generally the unique outcome if there is convergence.[13]

We focus on gradient learning by the slower firm. Gradient learning is a particularly naive strategy that requires minimal inputs. The slow firm does not take into account the actions or strategies of the faster rival, nor does it form beliefs about the economic environment. Gradient learning captures the fact that many firms simply adjust prices in the direction that increases static profits until profits are maximized. In particular, we assume that the slow firm begins with a best guess of its optimal price ($\hat{p}_b^0$). The firm updates its price based on the price in period t following

$$\hat{p}_b^{t+1} = \hat{p}_b^t + \lambda \frac{\partial \pi_b(\hat{p}_b^t)}{\partial \hat{p}_b^t}. \tag{15}$$

Gradient learning is also closely related to A/B testing in which firms use price experiments to determine whether to raise or lower prices, a commonly uses approach in online markets.[14] In a simultaneous game in which both firms use gradient learning, the strategies converges to the Betrand-Nash equilibrium as long as price adjustments are relatively smooth (λ is not too large).[15] We consider the case in which the slow firm employs gradient learning and ask whether the fast firm can implement a pricing strategy that results in supercompetitive profits.

In contrast to the previous sections, we assume here that the fast firm adopts a linear punishment strategy instead of a discontinuous trigger strategy. In particular, for a target price pair

[13] Fudenberg and Levine (2016) provide an overview of this literature. Asker et al. (2022) examine reinforcement learning and finds that synchronous reinforcement learning, in which firms can observe a rival's price and use this information to aide learning, converges to Bertrand-Nash. Asynchronous reinforcement learning, in which firms do not observe rival's price, can result in supercompetitive prices.

[14] Gradient learning is also similar to the simple "asynchronous" learning assumptions of Asker et al. (2022).

[15] See Anufriev et al. (2013).

(p_a^T, p_b^T), the faster firm chooses $p_a = p_a^T$ at the beginning of the period and then updates its price according to the linear pricing rule:

$$r_a = \begin{cases} p_a^T - \gamma(p_b^T - p_b) & \text{if } p_a^T - \gamma(p_b^T - p_b) \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{16}$$

Thus, the punishment depends on how far from the target price the slow firm deviates, and the degree of punishment is captured by γ.

We focus on linear strategies because the linearity helps ensure that experimentation by the slower firm will converge to desired price of the faster firm, for many different learning approaches of the rival firm (e.g., using Newton's method). In practice, pricing strategies that are linear in rivals price are common.

Throughout, we assume that the fast firm can fully commit to the linear pricing rule given by equation (16). This also implies the firm does not adjust its strategy to manipulate the rate of learning of the slower firm. Furthermore, we focus on the case with linear demand given by equation (10).

5.2 Coercive Linear Strategies with Simple Learning

We now solve for the pricing rule and equilibrium prices. The fast firm attempts to induce the target price vector (p_a^T, p_b^T). To constrain the slope of the reaction by the faster firm, we assume that the linear slope of the pricing rule passes through the point $(0, 0)$. Linear strategies of this form have the property that the faster firm's price changes in response to any non-negative price chosen by the slower firm. Using the pricing rule $r_a = p_a^T - \gamma^T(p_b^T - p_b)$, this implies $\gamma^T = \frac{p_a^T}{p_b^T}$.

In each period, the fast firm chooses its initial price p_a and target price for slow firm to maximize its own profit, such that the slow firm is maximizing static profit

$$\begin{aligned} &\max_{p_a, p_b^T} \left[\alpha p_a D_a(p_a, p_b) + (1-\alpha) \frac{p_a p_b}{p_b^T} D_a(\frac{p_a p_b}{p_b^T}, p_b) \right] \\ &s.t.\ \ p_b = \arg\max_{p_b'} \left[\alpha p_b' D_b(p_b', p_a) + (1-\alpha) p_b' D_b(p_b', \frac{p_a p_b'}{p_b^T}) \right] \end{aligned} \tag{17}$$

The slow firm's first-order condition is given by

$$\alpha \left[D_b(p_b, p_a) + p_b D_b^1(p_b, p_a) \right] + (1-\alpha) \left[D_b(p_b, \frac{p_a p_b}{p_b^T}) + p_b \frac{p_a}{p_b^T} D_b^2(p_b, \frac{p_a p_b}{p_b^T}) + p_b D_b^1(p_b, \frac{p_a p_b}{p_b^T}) \right] = 0 \tag{18}$$

The fast firm's pricing rule can be written in terms of the implicit function $p_b^*(p_a)$ that solves this first-order condition. The pricing-rule implies that

$$r_a = \frac{p_a p_b^*(p_a)}{p_b^T}. \tag{19}$$

In equilibrium, the fast firm sets the same price throughout the period so $p_a = r_a$. Let $p_a^*(p_b^T)$ be the fast firm price as a function of the target that solves $p_b^T = p_b^*(p_a)$ for p_a.

The fast firm then solves the following problem:

$$\max_{p_b^T} p_a^*(p_b^T) D_a(p_a^*(p_b^T), p_b^T). \tag{20}$$

The solution is given by the first-order condition

$$p_a^*(p_b^T)\left(\frac{\partial p_a^*(p_b^T)}{\partial p_b^T} D_a^1(p_a^*(p_b^T), p_b^T) + D_a^2(p_a^*(p_b^T), p_b^T)\right) + \frac{\partial p_a^*(p_b^T)}{\partial p_b^T} D_a(p_a^*(p_b^T), p_b^T) = 0 \tag{21}$$

This first-order condition provides the fast-firm optimal price for the slow firm, p_b^T. The optimal price for the fast firm is then $p_a^T = p_a^*(p_b^T)$. The solution reflect the fact that the fast firm chooses the target price for the slow firm knowing that the slow firm will maximize profit.

Proposition 2. *For the linear demand case, when the fast firm implements the optimal linear pricing rule, the slow firm profits are increasing in price for any $p_b < p_b^T$ and slow firm profits are decreasing in price for any $p_b > p_b^T$. Therefore, when firm b uses gradient learning it will always converge to the target price chosen by firm a.*

Proof. First, we solve explicitly for equilibrium target prices

$$p_a^T = -\frac{6d+2}{2\alpha d^2 + 4d + 1}$$

$$p_b^T = \frac{2 - 2(\alpha - 6)d(d+1)}{(2d+1)(2d(\alpha d + 2) + 1)}.$$

This yields the pricing rule

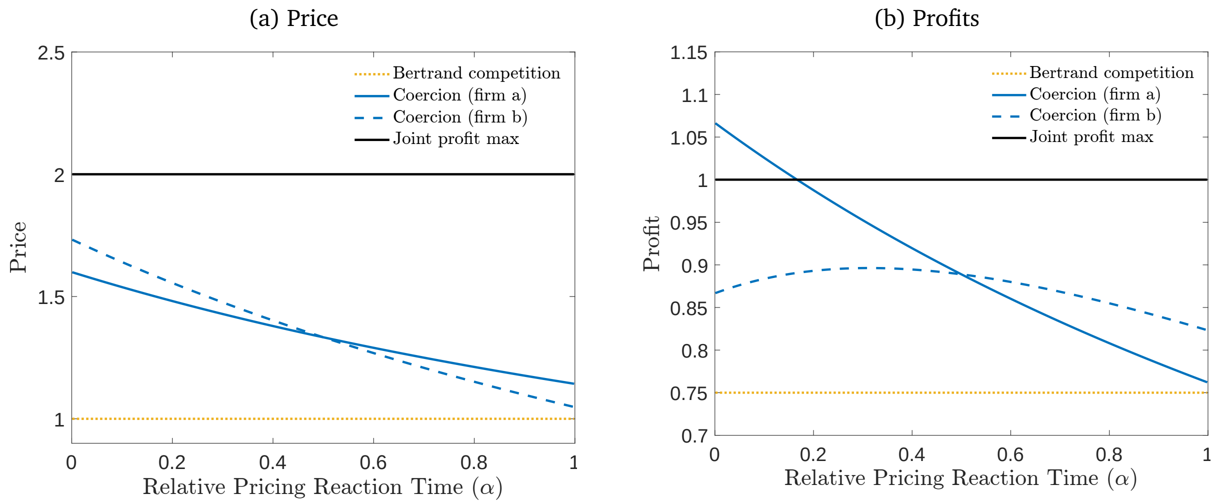
$$r_a(p_b) = -\frac{(2d+1)(3d+1)}{(\alpha - 6)d(d+1) - 1} p_b.$$

The slow firm chooses price p_b. The slow firm's profit function is given by $\tilde{\pi}_b(p_b) = \alpha \pi_b(p_b^*, p_a^T) + (1-\alpha)\pi_b(p_b^*, r_a(p_b))$. We solve for it explicitly:

$$\tilde{\pi}_b(p_b) = \frac{p_b(d(\alpha + 5\alpha d + 4) + 1)(2d(2(\alpha - 6) + \alpha d(2dp_b + \text{pb} + 2) + 4d(p_b - 3) + 3p_b) + p_b - 4)}{4((\alpha - 6)d(d+1) - 1)(2d(\alpha d + 2) + 1)}$$

For this function, it is the case that $\frac{\partial \tilde{\pi}_b(p_b)}{\partial p_b} > 0$ when $p_b > p_b^T$ and $\frac{\partial \tilde{\pi}_b(p_b)}{\partial p_b} < 0$ when $p_b < p_b^T$.

Figure 6: Coercive Strategies with Linear Punishment



Notes: Panel (a) shows prices and Panel (b) shows profits under Bertrand competition with simultaneous pricing, coercion 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, the firms simultaneously set prices at the beginning of the period and then the fast firm can react with lag α. Assumes $d = 1$ under linear demand given by equation (10).

Thus, gradient learning convergence to the optimum will yield the target prices. □

We depict the solution in Figure 6 using the linear demand system when $d = 1$. Panel (a) displays prices for different values of α. The fast firm prices are in the solid blue line. They are consistently higher than the Bertrand prices, and they increase with faster pricing (lower α). The fast firm's price is lower than the slow firm's price when the relative pricing reaction time is sufficiently small. These patterns are similar to the non-linear coercive strategies from the previous section (Figure 3). However, given the linear restriction on punishment, the fast firm cannot coerce its slow rival to set prices higher than the collusive price, indicating that the linearity of the strategies does limit the degree to which prices increase.

Panel (b) of Figure 6 displays the profits. Profits for the fast firm are declining in α, but profits for the slow firm are non-monotonic. With a fast enough reaction time, the fast firm can make higher profits than (its share of) the full collusion profits, even with the linear restriction. With $\alpha = 0$, the slow firm earns higher profits than in the case with coercive non-linear strategies (Figure 3).

6 Extension to N-Firm Oligopoly

Throughout the paper, we consider the case where a fast firm faces a single, slower rival. It is straightforward to extend the results to a more general setting in which a single fast firm faces $N - 1$ naive slower rivals that all have the same pricing frequency. Thus, the reaction time α

characterizes the reaction time that the faster firm has relative to each of the slower firms.

To extend the results, note that the profit function for any focal slow firm i can be written as

$$\pi_b(p_i, p_a, \{s_b\}) \tag{22}$$

where $\{s_b\}$ is the set of strategies of all other slower rivals. Under the assumption that the slower rivals are naive, all of the slower firm pricing decisions will be made independently holding fixed the strategies of rivals.

For a given price vector, we can define $\tilde{\pi}_i(p_i, p_a) = \pi_b(p_i, p_a, \{s_b\})$ for each slow firm and construct the relevant incentive compatibility conditions given the reaction time of the faster firm. Because the addition of these rivals does complicate the problem for the myopic firms, it is straightforward to modify the general problem from Section 4 to incorporate the target prices for additional firms and additional incentive compatibility constraints. As in our baseline model, maintaining a punishment function that is equal to the one-shot best response can yield unique equilibria. More generally, the sophisticated firm could potentially obtain even higher profits when it can specify an idiosyncratic reaction to individual deviations from each rival.

7 Conclusion

Asymmetries in pricing technology expand the set of equilibrium strategies that yield supra-competitive profits. A firm with faster pricing may incentivize a slower firm to set prices that maximize joint profits even when the slower firm is myopic, memoryless, or non-strategic. Thus, a firm with a pricing speed advantage can unilaterally—without the cooperation of its rivals—obtain identical outcomes to those obtained from collusion.

Firms with faster pricing may not want to induce the joint profit maximizing outcome. We characterize the set of coercive strategies that use punishment to maximize the faster firm's profits. This provides a different equilibrium, with prices that typically differ across firms even when profit functions are symmetric. The use of coercive strategies can be worse for consumers than joint profit maximization.

Overall, our results suggest a broader scope for firms to strategically increase prices. Sophisticated firms may be able to manipulate their rivals into setting prices above the competitive levels even when characteristics of the market would rule out traditional collusive strategies, such as short-termism or large differences in prices among similar firms. There is an opportunity for future research to examine the extent to which pricing strategies and algorithms used in practice may raise prices based on the features we identify here.

References

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

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

Aparicio, D., Z. Metzman, and R. Rigobon (2021). The pricing strategies of online grocery retailers. Technical report, National Bureau of Economic Research Working Paper.

Asker, J., C. Fershtman, and A. Pakes (2022). The impact of ai design on pricing. Technical report, Working Paper.

Assad, S., R. Clark, D. Ershov, and L. Xu (2022). Algorithmic Pricing and Competition: Empirical Evidence from the German Retail Gasoline Market. CESifo Working Paper 8521.

Bernheim, B. D. and M. D. Whinston (1990). Multimarket contact and collusive behavior. *The RAND Journal of Economics*, 1–26.

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

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

competitoor.com (2023). The importance of dynamic pricing. https://competitoor.com/pricing/the-importance-of-dynamic-pricing/.

dealhub.io (2023). Dynamic pricing. https://dealhub.io/glossary/dynamic-pricing/.

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.

Harrington, J. E. (2018). Developing Competition Law for Collusion by Autonomous Artificial Agents. *Journal of Competition Law & Economics 14*(3), 331–363.

Hortaçsu, A., A. Oery, and K. R. Williams (2022). Dynamic price competition: Theory and evidence from airline markets. Technical report, National Bureau of Economic Research.

Johnson, J., A. Rhodes, and M. R. Wildenbeest (2021). Platform Design When Sellers Use Pricing Algorithms. CEPR Discussion Paper No. DP15504.

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.

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.

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

Mouraviev, I. and P. Rey (2011). Collusion and leadership. *International Journal of Industrial Organization 29*(6), 705–717.

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

repricer.com (2023). Repricer.com, fastest amazon repricer. https://www.repricer.com/.

Scherer, F. M. (1980). *Industrial Market Structure and Economic Performance: 2d Ed*. Rand Mcnally College.

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

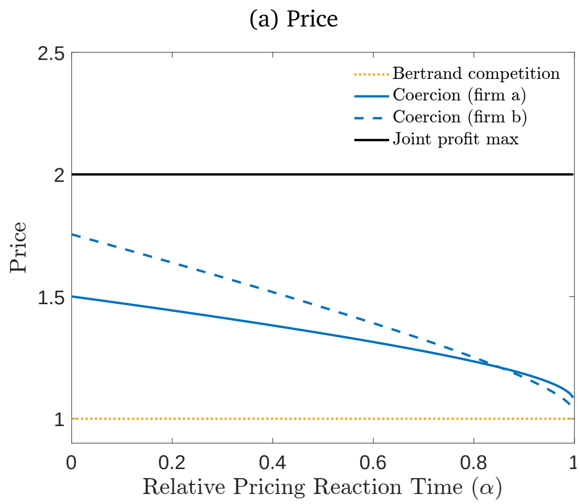
Tirole, J. (1988). *The Theory of Industrial Organization*. MIT Press, Cambridge, MA.

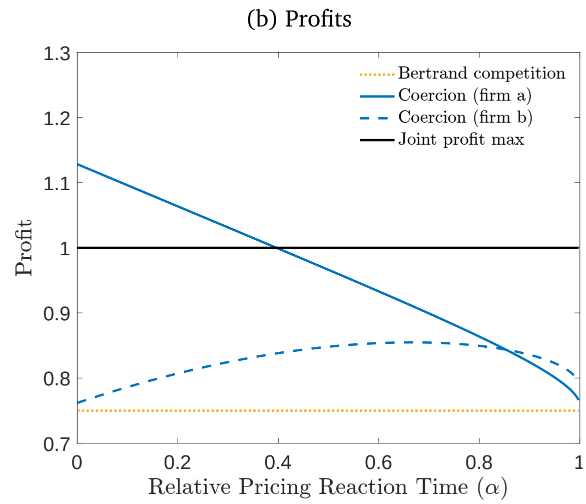
Waltman, L. and U. Kaymak (2008). Q-learning agents in a cournot oligopoly model. *Journal of Economic Dynamics and Control 32*(10), 3275–3293.

Appendix

A Additional Figures

Figure A-1: Prices and Profits with Coercive Trigger Strategies
With Partial Commitment ($\beta\gamma = 0.5$)

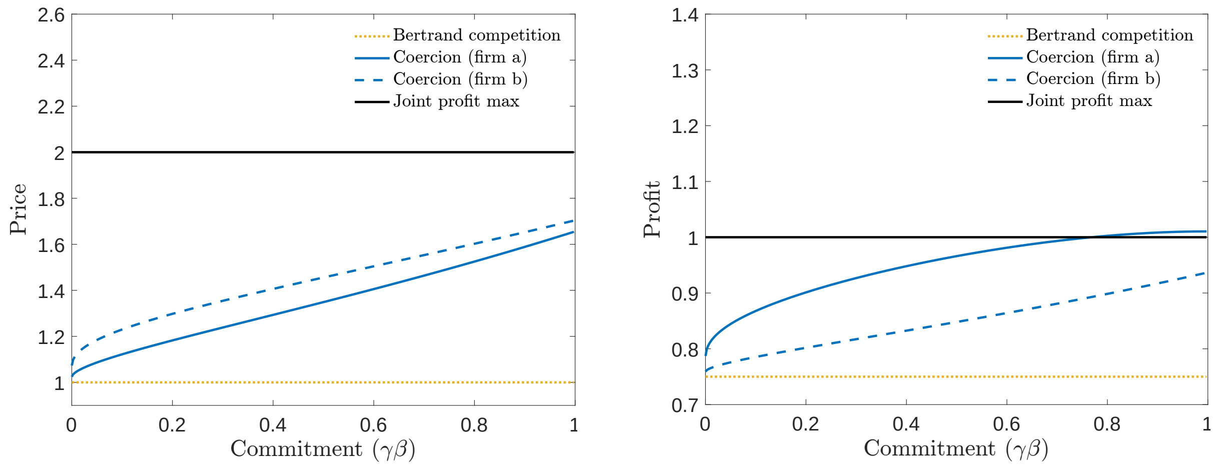




Notes: Panel (a) shows prices and panel (b) shows profits under Bertrand competition with simultaneous pricing, coercion in which the faster firm uses a trigger strategy to maximizes its own profits with partial commitment ($\beta\gamma = 0.5$), and joint profit maximization. The firms simultaneously set prices at the beginning of the period and then the fast firm can adjust price with reaction time α. Assumes $d = 1$ under linear demand given by equation (10).

Figure A-2: Prices and Profits with Coercive Trigger Strategies by Level of Commitment With Intermediate Reaction Time ($\alpha = 0.5$)

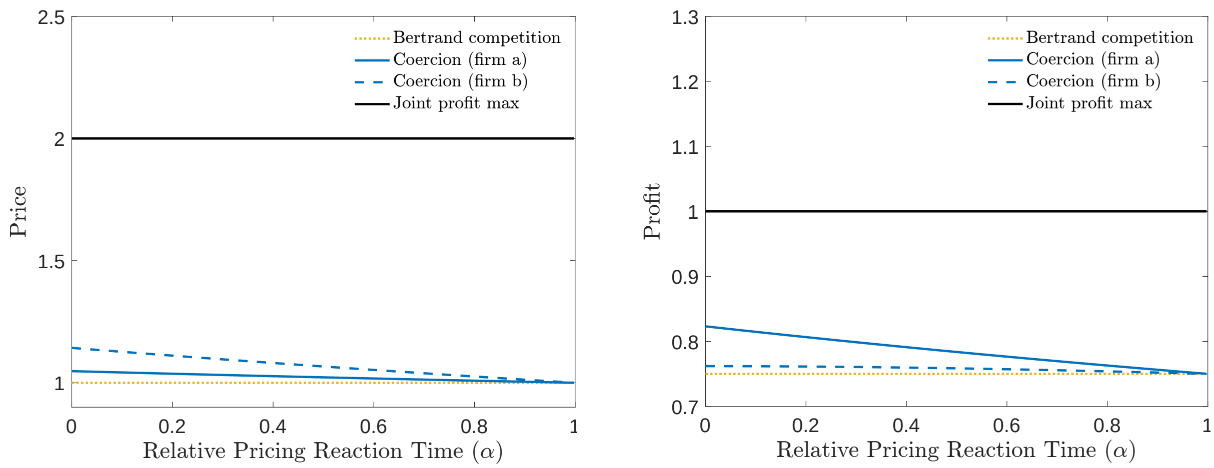
(a) Price (b) Profits



Notes: Panel (a) shows prices and panel (b) shows profits under Bertrand competition with simultaneous pricing, coercion in which the faster firm uses a trigger strategy to maximizes its own profits with $\alpha = 0.5$, and joint profit maximization. The firms simultaneously set prices at the beginning of the period and then the fast firm can adjust price with reaction time $\alpha = 0.5$. Assumes $d = 1$ under linear demand given by equation (10).

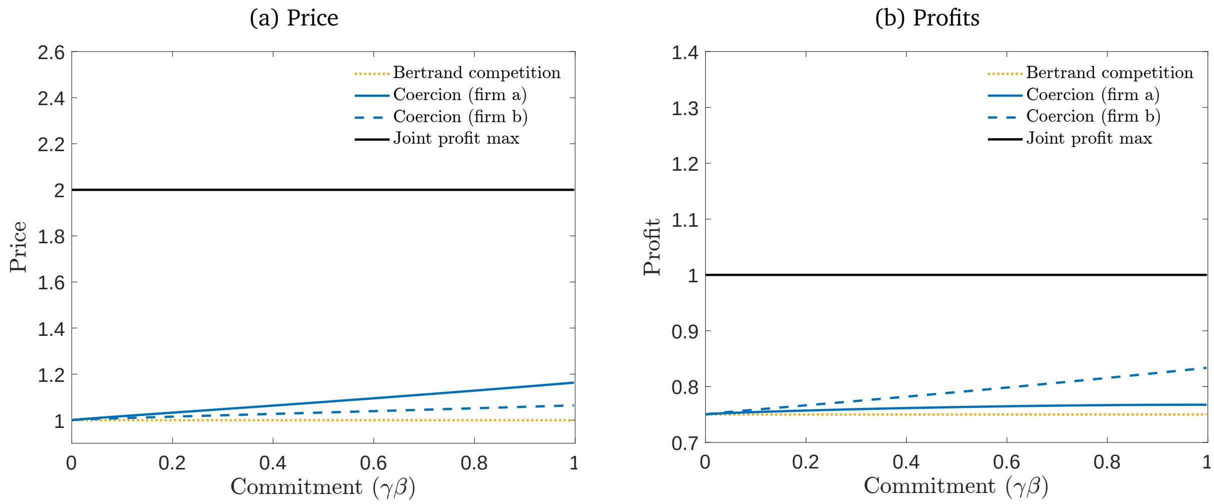
Figure A-3: Prices and Profits with Coercive Trigger Strategies With No Commitment ($\beta\gamma = 0$)

(a) Price (b) Profits



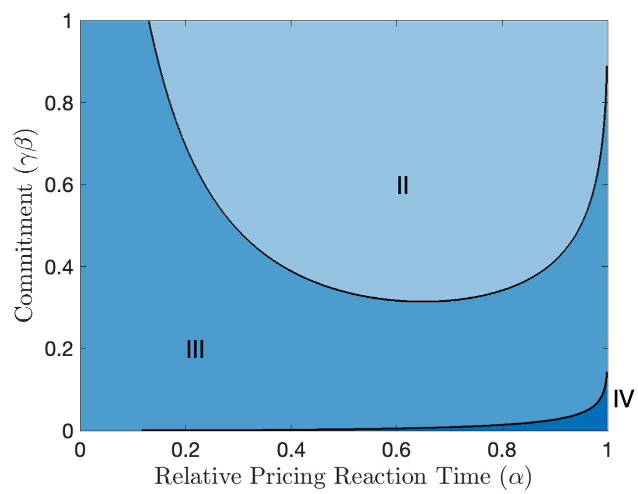
Notes: Panel (a) shows prices and panel (b) shows profits under Bertrand competition with simultaneous pricing, coercion in which the faster firm uses a trigger strategy to maximizes its own profits with no commitment ($\beta\gamma = 0$), and joint profit maximization. The firms simultaneously set prices at the beginning of the period and then the fast firm can adjust price with reaction time α. Assumes $d = 1$ under linear demand given by equation (10).

Figure A-4: Prices and Profits with Coercive Trigger Strategies by Level of Commitment With Simultaneous Timing ($\alpha = 1$)



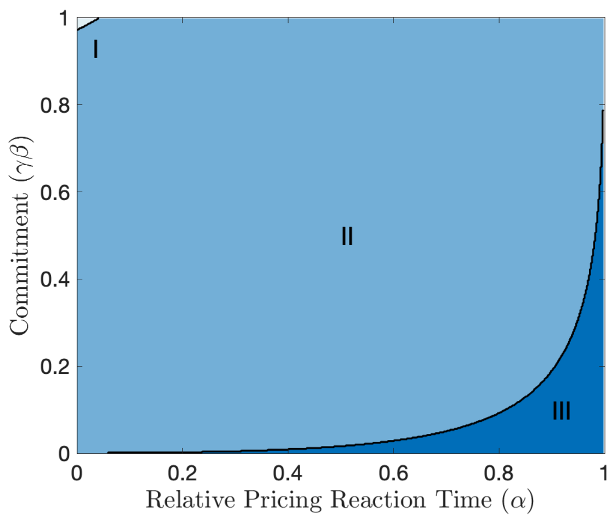
Notes: Panel (a) shows prices and panel (b) shows profits under Bertrand competition with simultaneous pricing, coercion in which the faster firm uses a trigger strategy to maximizes its own profits with $\alpha = 1$, and joint profit maximization. Assumes $d = 1$ under linear demand given by equation (10).

Figure A-5: Firm b Profits by Pricing Speed and Commitment



Notes: Shows profit regions of firm b for different values of reaction time (α) and commitment ($\beta\gamma$). Region II indicates profits greater than sequential follower, region III indicates profits greater than sequential leader, and region IV indicates profits greater than under Bertrand competition. Assumes $d = 1$ under linear demand given by equation (10).

Figure A-6: Consumer Surplus by Pricing Speed and Commitment



Notes: Shows profit regions of firm b for different values of reaction time (α) and commitment ($\beta\gamma$). Region I indicates consumer surplus less than under collusion, region II indicates consumer surplus less than the sequential equilibrium, and region III indicates consumer surplus less than under Bertrand competition. Assumes $d = 1$ under linear demand given by equation (10).