Supplemental Appendix to
Pricing with algorithms*

Rohit Lamba Sergey Zhuk

July 2026

forthcoming in American Economic Review: Insights

This note provides the details for the extensions described in Lamba and Zhuk [2026]. It has five substantive parts, followed by proofs. The first part introduces the extension to subgame perfect equilibrium and provides a detailed analysis of the two-price model. The second part discusses in some detail the relationship of the paper to Salcedo [2015]. The third part, which is the most substantial part of the appendix, delivers the proof of Theorem 2: the extension of the Markov perfect equilibrium analysis to a multiple-price grid. The fourth part relaxes Assumption 2 from the paper to include the pre-convergence part of payoffs into the calculations. The fifth part provides a conceptual foundation for Assumption 1 in the paper, that given our definition of algorithm, learning is cheap, since its cost is at most linear in time and length of the price grid.

1 Subgame perfect equilibria

The definition of an algorithm is the same, and the sellers still adjust their algorithms asynchronously; however, now the response (or choice) of each can depend not only on the current algorithm of the opponent, but on the overall history of algorithms chosen up to this point. We first describe a general procedure to calculate the set of subgame perfect equilibrium payoffs, and then deliver the theoretical lessons for low- and high-effective discount factors.

Suppose that one of the sellers is adjusting their algorithm and suppose that the current algorithm of their opponent is s. Define the set of all possible continuation payoffs (u, v) in a subgame perfect equilibrium depending on s (u corresponds to the seller adjusting their algorithm and v to the other one):

$$H(s) = \{(u, v) : \ (u, v) \text{ is possible in SPE starting with s}\} \tag{1}$$

There is a recursive way to construct these sets in the spirit of Abreu, Pearce, and Stacchetti [1990]. Suppose the seller adjusting their algorithm decides to respond to s with s', then the set of possible continuation payoffs is as follows:

$$H(s, s') = \left\{ (u, v) : \begin{array}{l} u = (1-\beta) \cdot \pi(s', s) + \beta \cdot v' \\ v = (1-\beta) \cdot \bar{\pi}(s, s') + \beta \cdot u' \\ (u', v') \in H(s') \end{array} \right\} \tag{2}$$

*Lamba: Cornell University, rohitlamba@cornell.edu; Zhuk: Saïd Business School, University of Oxford, Sergey.Zhuk@sbs.ox.ac.uk

$A\backslash B$	p_M	p_C
p_M	2, 2	0, 3
p_C	3, 0	1, 1

Table 1: Payoffs for SPE example.

The lowest possible continuation payoff u of the seller adjusting their algorithms is then:

$$u_{min}(s) = \max_{s'} \min_{(u,v)\in H(s,s')} u \tag{3}$$

Finally the set of possible SPE payoffs starting with s is:

$$H(s) = F(H)(s) = \bigcup_{s'} \left\{ (u,v) : \begin{array}{l} (u,v) \in H(s,s') \\ u \geq u_{min}(s) \end{array} \right\} \tag{4}$$

Any (u,w) with $u \geq u_{min}(s)$ could be sustained in an SPE since for any deviation there exists a subgame which generates a payoff no higher than u.

Now, to construct the sets of possible continuation payoffs $H(s)$, we can start with some arbitrary payoff sets that are guaranteed to contain $H(s)$, for example (here $\pi_{min} = \min_{p,q} \pi(p,q)$ and $\pi_{max} = \max_{p,q} \pi(p,q)$):

$$H_0(s) = [\pi_{min}, \pi_{max}] \times [\pi_{min}, \pi_{max}] \tag{5}$$

And, to calculate the sets $H(s)$, we can use the mapping F from equation (4) (which maps the set of payoffs $H = \{H(s)\ for\ all\ s\}$ onto itself) and then construct the sets of possible SPE payoffs as:

$$\begin{aligned} H_{n+1} &= F(H_n) \\ H(s) &= \lim_{n\to\infty} H_n(s) \end{aligned} \tag{6}$$

Note that for any n: $H_{n+1}(s) \subset H_n(s)$, thus, the limit exists. Moreover, the limit is exactly $H(s)$: for any $(u,v) \in \lim_{n\to\infty} H_n(s)$, we can construct a SPE by recursively calculating the possible continuation payoffs in the next periods (which always stay within $\lim_{n\to\infty} H_n(s)$ and thus can always be extended further).

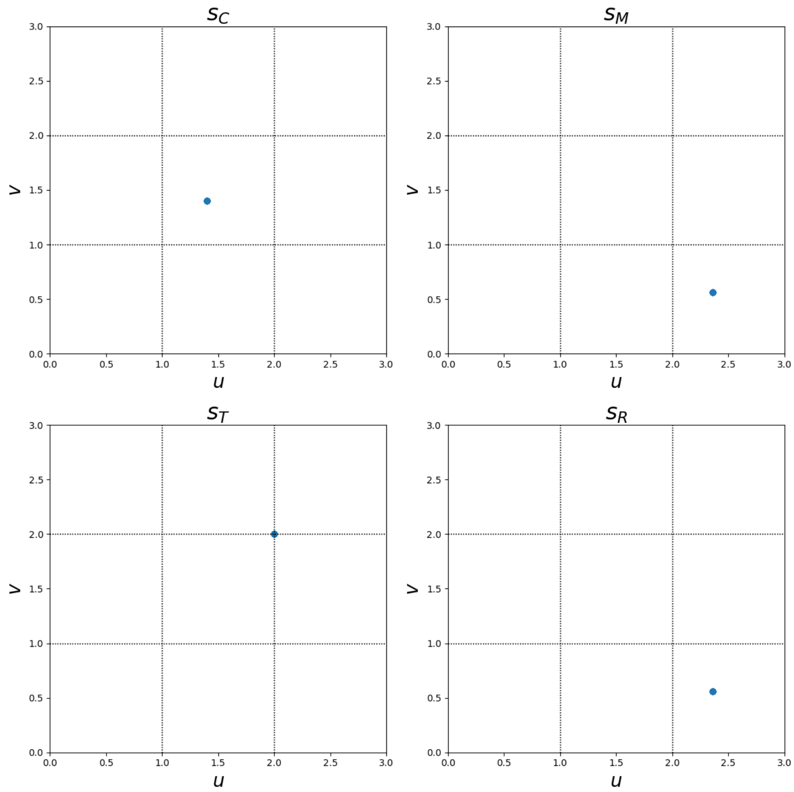
To illustrate the possible equilibria, consider the two-price set-up with payoffs for each customer from Table 1. For given payoffs we can construct the sets $H(s)$ numerically.[1] Figure 1 shows the sets of possible payoffs for each possible algorithm of the opponent for two values of the discount factor β. We make two points through these simulations.

Pure monopoly for low discounting. If the discount factor is low, then the only possible SPE leads to the monopoly outcome (p_M, p_M). In fact, our simulations suggest that the threshold on the effective discount factor does not have to be in the neighborhood of zero. As can be seen in Figure 1, this claim is valid for $\beta = 0.4$ for any starting strategy. The following result formalizes the claim with a sufficient threshold on β.

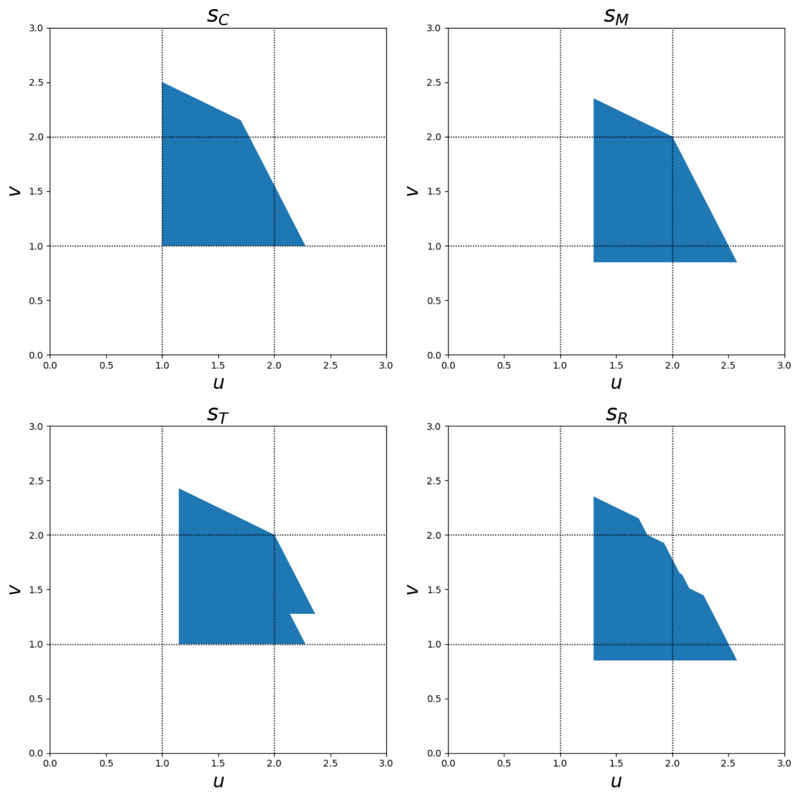
Theorem 3. *For sufficiently small β, in particular, if $\frac{\beta}{1-\beta} < \frac{min(\frac{1}{4},x,y)}{1+x+y}$, the only possible SPE is one in which both sellers respond to the current algorithm of the opponent as follows:*

$$s_R, s_M \longrightarrow s_C \longrightarrow s_T \longrightarrow s_T$$

[1] We do that by splitting the initial set $H_0(s)$ into a fine grid of squares (10000 by 10000). Then for each n the sets H_{n+1} contain only the squares for which at least one point (u,v) is feasible with a continuation pay-off from H_n. We also add a small margin to allow for mistakes in numerical calculations. Thus, the constructed sets are guaranteed to include $H(s)$, although they could be a bit larger.



(a) $\beta = 0.4$



(b) $\beta = 0.85$

Figure 1: Possible continuation payoffs in a subgame perfect equilibrium.

The equilibrium converges to the algorithm s_T for each seller and thus (p_M, p_M) as the outcome.

Proof. Let s denote the current algorithm of the opponent seller, and $SPE(s)$ the set of optimal responses (algorithms) possible in SPE for an algorithm s.

1. For sufficiently small β: $SPE(s_M) \subset \{s_C, s_R\}$.
   - The responding seller can get an immediate payoff of $1 + x$ (by choosing s_C or s_R) and would always do that for sufficiently small β.
2. For sufficiently small β: $SPE(s_T) = \{s_T\}$.
   - For sufficiently small β the seller would choose s_T or s_M to generate an immediate payoff of 1 (instead of $\pi(s_T, s_C) = 0$ or $\pi(s_T, s_R) = \frac{1}{2} + \frac{x-y}{4} < \frac{3}{4}$).
   - We can exclude s_M because s_M would generate the payoff of $-y$ (instead of 1 for s_T) after the opponent adjusts the algorithm next time.
3. For sufficiently small β: $SPE(s_C) = \{s_T\}$
   - For sufficiently small β it could be optimal to respond to s_C only with s_T or s_C.
   - We can exclude s_C because it would lead to the competitive payoff of 0 instead of the monopoly payoff of 1 after the opponent adjusts the algorithm next time.
4. For sufficiently small β: $SPE(s_R) = \{s_C\}$.
   - Only s_R or s_C could be optimal for s_R for sufficiently small β.
   - We can exclude s_R because it generates the subsequent payoff of $-y$ (instead of 0) after the opponent adjusts the algorithm next time.
5. For sufficiently small β: $SPE(s_M) = \{s_C\}$
   - s_R can be excluded since it generates the subsequent payoff of $-y$ (instead of 0 for s_C) after the opponent adjusts the algorithm next time.

Note that the highest possible continuation payoff is $1+x$ and the lowest is $-y$, which gives us the following sufficient conditions for each of the points that are made above.

1. $(1-\beta) \cdot (1+x) + \beta \cdot (-y) > (1-\beta) \cdot 1 + \beta \cdot (1+x)$
2. $(1-\beta) \cdot 1 + \beta \cdot (-y) > (1-\beta) \cdot 0 + \beta \cdot (1+x)$
   $(1-\beta) \cdot 1 + \beta \cdot (-y) > (1-\beta) \cdot \left(\frac{1}{2} + \frac{x-y}{4}\right) + \beta \cdot (1+x)$
   $(1-\beta) \cdot 1 + \beta \cdot (-y) > (1-\beta) \cdot (-y) + \beta \cdot (1+x)$
3. $(1-\beta) \cdot 0 + \beta \cdot (-y) > (1-\beta) \cdot (-y) + \beta \cdot (1+x)$
   $(1-\beta) \cdot 1 + \beta \cdot (-y) > (1-\beta) \cdot 0 + \beta \cdot (1+x)$
4. $(1-\beta) \cdot (1+x) + \beta \cdot (-y) > (1-\beta) \cdot (-y) + \beta \cdot (1+x)$
   $(1-\beta) \cdot 0 + \beta \cdot (-y) > (1-\beta) \cdot (-y) + \beta \cdot (1+x)$
5. $(1-\beta) \cdot 0 + \beta \cdot (-y) > (1-\beta) \cdot (-y) + \beta \cdot (1+x)$

Combining these conditions together, we get the overall sufficient condition for the unique SPE described in the theorem.

$$\frac{\beta}{1-\beta} < \frac{\min\left(\frac{1}{4}, x, y\right)}{1+x+y} \implies \frac{\beta}{1-\beta} < \frac{\min\left(\frac{1}{2}+\frac{y-x}{4}, x, y\right)}{1+x+y}$$

□

This confirms the result asserted in Salcedo [2015] that as the effective discount factor converges to zero, repeated play of the monopoly outcome emerges as the unique SPE. In contrast to Salcedo [2015], the result is clearer to see here for the two-price model and a sufficiently sharp threshold on the smallness of effective discounting can also be provided.

This result may be counterintuitive for it breaks the standard intuition that as the horizon of players gets shorter, the unique static Nash (in this case competitive pricing (p_C, p_C)) emerges as the only subgame perfect equilibrium. In contrast, here monopoly pricing emerges as the unique equilibrium outcome.

High discounting leads to large and complex equilibrium payoff sets. As is widely understood from folk theorems, for high discount factors, equilibria expand to the set of all payoffs above a certain minimal threshold. This also holds for the asynchronous repeated game we study.[2] However, the actual strategies required to sustain these payoffs are complex. Figure 1 shows the set of equilibrium payoffs for $\beta = 0.85$.

Consider, for example, the competitive equilibrium in which each seller always chooses s_C, which would generate continuation payoffs of $(1, 1)$ for the payoff matrix from Table 1. Suppose that one of the sellers deviates to s_T in one of the rounds, and the subsequent continuation payoffs are $(u, v) \in H(s_T)$. In order for this deviation to be not profitable, v should be equal to 1. But note that u is non-trivially different from 1 ($u \geq 1.15$), thus, we should expect a complicated sequence of price pairs and corresponding algorithms that would generate such payoffs. For example, consider the lowest point ($u = 1.15$ and $v = 1$). For this point, we can construct the possible continuation sequences of algorithms that could generate such payoffs. We do this by choosing the algorithm s' such that $(u, v) \in H(s_T, s')$ and then doing this recursively. In the end, we get the following sequence of algorithms (" $\times$ 13" means s_C is repeated 13 times):

$$s_T, s_C \times 13, s_M, s_C, s_M, s_C \times 2, s_M \times 10, \ldots, \text{[3]} \tag{7}$$

The above sequence of responses requires some very specific beliefs about how the opponent should behave in equilibrium while choosing algorithms, which we find unrealistic. This is the reason why we decided to focus only on Markov perfect equilibria, which restrict the complexity of the possible strategies.[4]

To be sure, the high discounting case is the more realistic one. So, the conceptual way to digest these two theoretical arguments is that even though a high discount factor with SPE nudges results towards the folk theorem, the fact that with low discount factors only the collusive outcomes survive underscores the salience of the collusive outcome in this algorithmic pricing problem, even when other, more complex, and arguably unrealistic, equilibria are possible. In fact, one can go further and conjecture that simple tit-for-tat-type algorithms are particularly attractive in enabling collusion when pricing is done by competing codes.

[2] Establishing this formally is beyond the scope here, but it is informally confirmed by computations.

[3] Here the possible continuation algorithms were always reviewed in the following order: s_C, s_M, s_T, s_R. Thus, if there were several possibilities with which algorithm could continue, the earliest one in the above sequence was always selected. This also means that other sequences of algorithms are certainly possible. However, all such sequences have similar irregular structure (we also constructed all possible sequences up to 15 elements and manually checked them).

[4] In terms of the repeated games literature if we allowed for public randomization, then a lot of the complexity in achieving certain payoffs can be pushed into the randomization devices. However, it is not clear how to think about public randomization when it comes to modeling algorithms.

2 A comparison to Salcedo (2015)

The above extension to subgame perfect equilibrium provides a good segue to compare our analysis with Salcedo [2015]. It also looks at a duopoly model of price-setting with probabilistic demand and algorithms as strategies, but uses subgame perfectness as the equilibrium criterion. The basic idea of using such a set up to explore algorithmic collusion is elegant. We depart from its analysis in several ways.

First, we model the competition in algorithms differently. In Salcedo [2015], whenever a customer arrives, the price is set by a complex automaton. In our model customer arrivals are unobservable and the algorithms react only to prices set by the competitor. We believe this assumption lends realism to how algorithms actually interact, especially when we think about online sellers. In addition, it allows us to separate the frequency of customer arrivals from the frequency of possible price changes.

Further, we assume that algorithms can react fast to possible price changes by the competitor, again a reasonable assumption. This allows us to replace the complicated continuous time game with a simpler asynchronous repeated game in which payoffs in each period are determined by the convergent prices for the current algorithms. This, in turn, allows us to derive sharper results, which would be much more difficult to obtain otherwise.

The second key difference is that in Salcedo [2015] the main result about the inevitability of monopoly outcomes is stated only for the limiting case in which the frequency of algorithmic adjustments (μ in our model) converges to 0.[5] Then the effective discount factor $\beta = \frac{\mu}{\mu+r}$ also converges to 0. This limits the dynamic dimension of the strategic interaction between sellers, for they essentially behave "myopically," best responding only to the current algorithm of the opponent, not taking into account how the opponent would react to their choice.

The case of $\beta \approx 0$ is also unrealistic. Suppose we fix the discount factor at $r = 5\%$ per year. Then to generate $\beta = 0.5$, which is not even that low, we would need $\mu = 0.05$. In plain speak, this means the sellers have the opportunity to revise their algorithms, on average, once in 20 years. If $\beta = 0.1$, then it is once in 180 years. More realistically, suppose we fix $\mu = 5$, which means each seller has on average five opportunities per year to revise their algorithm. Then, $\beta \approx 0.99$. For $\mu = 10$, $\beta \approx 0.995$.

When β is high, as we show in Section 1 above, the monopoly outcome is no longer inevitable in subgame perfect equilibria, and a wide range of outcomes are possible (as is typical for repeated games). Nevertheless, we are able to show that some collusion is inevitable and extreme collusion is very likely for Markov perfect equilibria. Thus, we can conclude that even for high β's, which we argue are more realistic, algorithms will still facilitate collusion.

3 Multiple price grid

3.1 Updated model, main result, and some intuition

We introduced the multiple price grid setup in Section V of the main paper, Lamba and Zhuk [2026], along with the appropriate assumptions. Here is a brief recap: Suppose that prices belong to a discrete set of equally spaced numbers between the competitive p_C and monopoly prices p_M:

$$p_C = p_1 < \cdots < p_K = p_M. \tag{8}$$

[5] Salcedo [2015] also states a result as the revision rate goes to infinity. This means that sellers are changing algorithms faster than consumer arrivals, so before any payoffs accrue, and in the limit even faster than price changes, which is unrealistic.

Fix $\Delta = p_i - p_{i-1}$ for $i = 2, \ldots, K$ to be the difference between two consecutive prices. In a slight abuse of notation, we denote:

$$\begin{aligned} p_{M-\Delta} &= p_M - \Delta = p_{K-1} \\ p_{M-2\Delta} &= p_M - 2 \cdot \Delta = p_{K-2} \\ &\ldots \\ p_{C+\Delta} &= p_C + \Delta = p_2 \end{aligned} \tag{9}$$

Assumption 0′ places some minimal structure on payoffs, and Assumption 3 ruled out cycles to keep the analysis tractable, noting that we had allowed for cycles in the two-price model and the main result continued to hold. The main result with the multiple-price grid is stated as follows.

Theorem 2. *Suppose $s_0, s_1, s_2, \ldots$ is a path of algorithms in a Markov equilibrium. Then:*

1. *payoff of both sellers is higher than the competitive outcome:*

$$\forall\, k \geqslant 1, \min\{u^i_k, v^i_{k+1}\} \geqslant \pi(p_C, p_C)\ \forall\, i \in \{A, B\}, \textit{ and}$$

2. *payoff of at least one seller is eventually bounded below approximately by the monopoly outcome:*

$$\textit{for } i = A \textit{ or } i = B, \exists\, k \textit{ s.t. } \min_{m \geqslant k}\{u^i_m, v^i_{m+1}\} \geqslant (1-\beta) \cdot \pi(p_{M-\Delta}, p_{M-\Delta}) + \beta \cdot \pi(p_M, p_M).$$

While the first part is straightforward to absorb, the second one can be interpreted as follows: If the price grid is dense enough or the effective discount factor is high enough, then in any Markov perfect equilibrium, the payoff of at least one seller is bounded from below by the monopoly profit. Both point to the exercise of some market power by the sellers and taken together they establish this market power is significant. Notice, also, that even for the weakest case of $\beta \approx 0$ and only 3 possible prices in the grid the lower bound is at least $\pi\left(\frac{p_C+p_M}{2}, \frac{p_C+p_M}{2}\right)$ which is non-trivially higher than the competitive payoff.[6]

Notation: Define $\xi(s^{-i}, s^i) = (p^{-i}, p^i)$ as the price pair to which the algorithms would converge; it would be a price pair consistent with these two algorithms and since i was the last to adjust their algorithm, amongst all such price pairs it will be the one most preferred by i. And to simplify notation and reduce superscripts, in what follows, we write the profit function in the discussion and proof as $\pi(p, q)$, where p is the own price and q is the price of the opponent.

Intuition: The main characteristic that expands the set of equilibria is the existence of *unequal* equilibria, in which one seller makes a profit even higher than the monopoly outcome, $\pi(p_M, p_M)$. Here is a simple enough intuition: Suppose (p, q) is a pair of prices such that payoffs of the sellers at these prices $(\pi(p, q), \pi(q, p))$ are on the convex hull of the Pareto Frontier and in addition that

$$\pi(p_C, p_C) < \pi(q, p) < \pi(p_M, p_M) < \pi(p, q). \tag{10}$$

So, the payoff of one seller is above the monopoly outcome and the other is between the monopoly and competitive outcome. Now consider the following algorithms

$$\begin{aligned} s^i &= \{q \to p, x \to p_C \; for \; x \neq q\} \\ s^{-i} &= \{p \to q, x \to p_C \; for \; x \neq p\} \end{aligned} \tag{11}$$

[6]The reason why the lower bound is not exactly the monopoly payoff, is that for general payoff matrices with multiple prices there could also exist alternating price equilibria (similar to "alternating prices" equilibria from Theorem 1 in Lamba and Zhuk [2026]). We believe that for specific payoff functions (e.g., "linear" or "discrete choice") such equilibria might also be ruled out and the lower bound could be improved.

This is a grim trigger sort of strategy that aims to sustain the choice of (p, q) in equilibrium. Interestingly, such unequal equilibria can be constructed for an arbitrary discount factor β. If seller $-i$ believes that for any of their responses seller i will keep their current strategy s^i, then they also do not have any incentives to deviate from s^{-i}. And for seller i it would indeed be optimal to do that if $\pi(p, q)$ is sufficiently high. To illustrate these ideas, we explicitly construct such an equilibrium with 5 different prices (see Section 3.4). As in repeated games, it is, however, easier to operate in the payoff space and characterize key properties of the set of equilibrium payoffs, which is what we do in this section.

3.2 Proof outline of Theorem 2

The proof consists of several parts. In sections 3.2.1-3.2.3 we develop the necessary techniques. In 3.2.1 we show that the state space can be contracted from the set of algorithms to the matrix of price pairs, which significantly reduces the complexity of the analysis. Then in 3.2.2 we derive the conditions that the continuation payoff matrices should satisfy in order for them to correspond to an equilibrium. Section 3.2.3 then defines a notion of "best" and "second best" algorithms that each seller could have whenever their competitor is adjusting their algorithm.

Finally, we prove the theorem in section 3.2.4. We consider an arbitrary equilibrium path, and show that the expected continuation payoffs for both sellers tend to increase on this path, and if not, the corresponding seller can guarantee at worst the "second best" continuation payoff. Lemma 4 then shows that such payoff is above the limit stated in Theorem 2 for at least one of the sellers.

3.2.1 State space reduction

Whenever a seller responds with an algorithm that leads to a price pair (p, q) they could always choose the best algorithm consistent with this price pair (that maps q into p). So whenever we observe a convergent price pair (p, q) after seller i adjusted their algorithm, we should observe the same algorithm for seller i. Thus, the price pair (p, q) becomes essentially the state variable, and since there are much fewer price pairs than possible algorithms this allows us to significantly reduce the state space.

We can define

$$v^i(p, q) = \max_{s:q \to p} v^i(s) \tag{12}$$

$$s^i(p, q) = \arg\max_{s:q \to p} v^i(s) \tag{13}$$

$$u^{-i}(q, p) = u^{-i}(s^i(p, q)) \tag{14}$$

If we know the payoff matrices $v^i(p, q)$ and $v^{-i}(q, p)$ then for any algorithm s of seller $-i$ we can calculate the highest possible continuation payoff that seller i could get:

$$u^i(s) = \max_p (1 - \beta) \cdot \pi(p, s(p)) + \beta \cdot v^i(p, s(p)) \tag{15}$$

Even though this equation does not tell us with which exact algorithms seller i should respond, it pins down the highest payoff they can ensure while responding to an algorithm s by $-i$.

3.2.2 Payoff matrices in equilibrium

In this section we will derive the conditions that matrices $v^i(p, q)$ and $u^{-i}(q, p)$ for $i = A, B$ should satisfy if they correspond to payoffs in a certain Markov perfect equilibrium. Also, we will show how to construct the best response to any algorithm if the values of the payoff matrices are known.

First, let us define the worst possible algorithm for seller i to respond to, and the lowest continuation payoff that they can get when responding to some algorithm:

Definition 1. *The worst algorithm of seller $-i$ for seller i is defined as:*

$$\underline{s}^{-i}(p) = \arg\min_q (1-\beta)\cdot\pi(p,q) + \beta\cdot v^i(p,q) \tag{16}$$

Definition 2. *Then the lowest possible continuation payoff of i when i is adjusting their algorithm is:*

$$\underline{u}^i = \max_p \left[\min_q (1-\beta)\cdot\pi(p,q) + \beta\cdot v^i(p,q)\right] \tag{17}$$

From equation (15) it follows that for any algorithm s:

$$u^i(s) \geq \underline{u}^i = u^i(\underline{s}^{-i}) \tag{18}$$

Now let's calculate $v^i(p,q)$. Suppose i chooses an algorithm such that the next pair of prices (after $-i$ adjusts their algorithm) is (p',q'). Then for this price pair we can define the "before" continuation payoffs $V^i(p',q')$ and $U^{-i}(q',p')$, which include this price pair and assume optimal behavior afterwards:

Definition 3. *"Before" continuation payoffs for price pair (p',q') are defined as*

$$\begin{aligned} V^i(p',q') &= (1-\beta)\cdot\pi(p',q') + \beta\cdot u^i(p',q') \\ U^{-i}(q',p') &= (1-\beta)\cdot\pi(q',p') + \beta\cdot v^{-i}(q',p') \end{aligned} \tag{19}$$

Then, if price pair (p',q') follows (p,q) the continuation payoffs are linked as:

$$v^i(p,q) = V^i(p',q') \qquad u^{-i}(q,p) = U^{-i}(q',p') \tag{20}$$

Note that the values of the "before" continuation payoffs $V^i(p',q')$ and $U^{-i}(q',p')$ are directly determined by the "after" continuation payoff matrices $u^i(p',q')$ and $v^{-i}(q',p')$. Finally, the following proposition links the "before" and "after" continuation payoffs, thus establishing the recursive relationship that the continuation payoff matrices $u^i(p',q')$ should satisfy.

Proposition 1. *For any price pair (p,q):*

$$v^i(p,q) = \max V^i(p',q') \qquad s.t. \quad \begin{aligned} &U^{-i}(q',p') \geq \underline{u}^{-i} \\ &U^{-i}(q',p') \geq U^{-i}(q,p) \\ &\textit{either},\ (p',q') = (p,q),\ \textit{or},\ q' \neq q \end{aligned} \tag{21}$$

Moreover, if the price pair (p',q') maximizes the above expression then $u^{-i}(q,p) = U^{-i}(q',p')$.

The idea of the proof is as follows. If price pair (p',q') goes after (p,q) then all conditions stated in equation (21) should be satisfied (the payoff of seller $-i$ can never be below $\underline{u}^{-i}$ and it cannot be below $U^{-i}(q,p)$ since $-i$ can keep the strategy from the previous period). In addition (see the proof of Proposition 1 below), by using the following algorithm:

$$s_*^i(p,q) = \{q \to p, q' \to p', x \to \underline{s}^i(x)\ for\ x \neq q\ and\ x \neq q'\} \tag{22}$$

seller i can make sure that the subsequent prices will actually be (p', q').

Proposition 1 gives us a powerful tool showing which price transitions are possible in an equilibrium. In addition, it gives us precise conditions that payoff matrices $v^i(p,q)$ and $u^{-i}(q,p)$ should satisfy in order for them to correspond to an equilibrium.

3.2.3 "First best" and "Second best" continuation payoffs

The final piece of terminology that we would need in order to prove Theorem 2 is the first best and the second best continuation payoffs and the corresponding algorithms that would allow sellers to achieve such payoffs.

Definition 4. *The "best" continuation payoff of i when $-i$ is adjusting their algorithm is defined as:*

$$v_*^i = \max V^i(p,q) \qquad s.t.\, U^{-i}(q,p) \geq \underline{u}^{-i} \tag{23}$$

Also, if (p_*^i, q_*^i) maximizes the above expression, then we define the best continuation algorithm of i as[7]*:*

$$s_*^i = \{q_*^i \to p_*^i, x \to \underline{s}^i(x) \; for \; x \neq q_*^i\} \tag{24}$$

Definition 5. *If (p_*^i, q_*^i) are the "first best" prices from definition 4 then we define the "second best" continuation payoff of seller i when $-i$ is adjusting their algorithm as:*

$$v_{**}^i = \max V^i(p,q) \qquad s.t. \quad \begin{array}{l} U^{-i}(q,p) \geq \underline{u}^{-i} \\ q \neq q_*^i \end{array} \tag{25}$$

Also if (p_{**}^i, q_{**}^i) maximizes the above expression (25) then we can define the "second best" continuation algorithm of i as:

$$s_{**}^i = \{q_{**}^i \to p_{**}^i, x \to \underline{s}^i(x) \; for \; x \neq q_{**}^i\} \tag{26}$$

Note that seller i cannot get a continuation payoff $v^i(s)$ higher than v_*^i for any s. Moreover, the payoff v_*^i is exactly achievable for $s = s_*^i$ (if seller $-i$ does not continue with prices (p_*^i, q_*^i) their payoff is at most $\underline{u}^{-i}$). The second best payoff v_{**}^i corresponds to the case when seller i cannot use the opponent's price q_*^i from the first best pair (p_*^i, q_*^i).[8]

3.2.4 Evolution of equilibrium continuation payoffs

Consider some arbitrary equilibrium sequence of algorithms $s_0, s_1, s_2, \ldots$ Note that "even" algorithms in the sequence $s_0, s_2, \ldots$ correspond to one seller, and "odd" algorithms to the other. Below when we write expressions like $u^i(s_k)$ we assume that i and k agree with each other (e.g. in this case it means that s_k is the algorithm of seller i).

Lemma 1 shows that each seller can guarantee that their continuation payoffs stay above the competitive payoff $\pi(p_C, p_C)$, which proves the first part of Theorem 2.

Lemma 1. $u^i(s_k) \geq \pi(p_C, p_C)$ *for* $k \geq 0$ *and* $v^{-i}(s_k) \geq \pi(p_C, p_C)$ *for* $k \geq 1$.

[7] In Definitions 4 and 5 if several feasible price pairs maximize $V^i(p,q)$ then we select the one maximizing $U^{-i}(q,p)$. As a general rule, as is standard, if a seller adjusting their algorithm is indifferent between two options (in terms of continuation payoffs), we assume that the seller would choose the one that is preferred by their opponent.

[8] If none of the price pairs satisfy the stated conditions for the "second-best" (this means seller $-i$ always ends up with price q_* after adjusting their algorithm), it is fairly easy to show that the continuation payoffs for at least one of the sellers in any equilibrium sequence will exceed $\pi(p_M, p_M)$. See Section 3.3 for details. So, in what follows we will assume that the second-best always exists.

We can also show that the continuation payoffs tend to increase on the equilibrium path (and if not, the corresponding seller can guarantee at least the second-best continuation payoff).

Lemma 2. *For any k and corresponding i:*

1. $u^{-i}(s_{k+1}) \geq v^{-i}(s_k)$
2. *Either* $v^i(s_{k+1}) \geq u^i(s_k)$ *or* $v^i(s_{k+1}) = v^i_*$ *or* $v^i(s_{k+1}) = v^i_{**}$*. Moreover, the last case is only possible if* $\xi(s_k, s_{k+1}) = (q^i_*, x)$ *for some x where q^i_* is from Definition 4.*

Suppose the continuation payoffs of one seller become constant, then we can show that the continuation payoffs of the other seller will also become constant. Moreover, at least one of these constant payoffs should be above the monopoly profit.

Lemma 3. *If for some i and k:*

$$u^i(s_k) = v^i(s_{k+1}) = u^i(s_{k+2}) = \dots \tag{27}$$

then for some m

$$u^{-i}(s_m) = v^{-i}(s_{m+1}) = u^{-i}(s_{m+2}) = \dots \tag{28}$$

Moreover, either $u^i(s_k) \geq \pi(p_M, p_M)$ *or* $u^{-i}(s_m) \geq \pi(p_M, p_M)$.

Finally, if continuation payoffs for both sellers never become constant then eventually both of them will be above the corresponding second-best payoffs v^i_{**}. Thus, to prove theorem 2 we only need to show that at least one of these second-best payoffs is above the bound stated in the theorem.

Lemma 4.

$$\max\left(v^i_{**}, v^{-i}_{**}\right) \geq (1-\beta)\cdot\pi(p_{M-\Delta}, p_{M-\Delta}) + \beta\cdot\pi(p_M, p_M) \tag{29}$$

3.3 Non-existent "Second best" price pair

In this section we consider a special case, in which the "second best" price pair from Definition 5 does not exist for at least one of the sellers. We can show that in this case for an arbitrary equilibrium sequence $s_0, s_1, s_2, \dots$ the continuation payoffs for at least one of the sellers will exceed the monopoly payoff $\pi(p_M, p_M)$. The steps below prove this result. For the "first best" pair we use the same notation equation (54), (p_*, q_*) and $(\dot{q}, \dot{p})$, as in the proof of Lemma 4.

1. Suppose, for example, $U^{-i}(q, p) < \underline{u}^{-i}$ for any $p, q \neq q_*$. This implies $U^{-i}(q_*, p) \geq \underline{u}^{-i}$ for any p (seller i can construct an algorithm which would lead to prices (q_*, p)).
2. As a result, Lemma 2 for seller i means either $v^i(s_{k+1}) \geq u^i(s_k)$ or $v^i(s_{k+1}) = v^i_*$. Thus, the continuation payoffs for seller i in the equilibrium sequence either become constant (we can use Lemma 3) or reach v^i_*.
3. If $q_* = p_M$ then $v^i_* \geq V^i(p_M, p_M) \geq \pi(p_M, p_M)$. Thus, assume $q_* \neq p_M$.
4. If $\dot{p} \neq p_M$ then $v^{-i}_{**} \geq V^{-i}(p_M, p_M) \geq \pi(p_M, p_M)$. Thus, assume $\dot{p} = p_M$. (Since $U^i(p_M, p_M) \geq \pi(p_M, p_M)$, the "second best" price pair necessarily exists for seller $-i$).
5. Since $U^{-i}(\dot{q}, \dot{p}) = (1-\beta)\cdot\pi(\dot{q}, \dot{p}) + \beta\cdot v^{-i}_* \geq u^{-i}(\{x \to p_C\}) \geq \underline{u}^{-i}$ then $\dot{q} = q_*$.
6. Finally for any p: $U^{-i}(q_*, p) = (1-\beta)\cdot\pi(q_*, p) + \beta\cdot v^{-i}(q_*, p)$. Then, either $p = \dot{p}$ and $v^{-i}(q_*, p) = v^{-i}_*$ or $p \neq \dot{p}$. This together with Lemma 2 implies that in any equilibrium sequence either continuation payoffs for seller $-i$ converge to a constant or they become above $v^{-i}_* \geq V^{-i}(p_M, p_M) \geq \pi(p_M, p_M)$.

$A\backslash B$	8	7	6	5	4
8	2.95, 2.95	2.41, 3.39	1.86, 3.61	1.36, 3.54	0.95, 3.19
7	3.39, 2.41	2.87, 2.87	2.29, 3.16	1.74, 3.21	1.25, 2.97
6	3.61, 1.86	3.16, 2.29	2.64, 2.64	2.08, 2.79	1.55, 2.68
5	3.54, 1.36	3.21, 1.74	2.79, 2.08	2.30, 2.30	1.80, 2.32
4	3.19, 0.95	2.97, 1.25	2.68, 1.55	2.32, 1.80	1.90, 1.90

Table 2: Expected payoffs from one customer in an unequal equilibrium.

Seller A

$A\backslash B$	8	7	6	5	4
8	8, 8	8, 7	8, 6	8, 6	8, 7
7	7, 8	8, 8	8, 7	8, 7	8, 7
6	5, 4	7, 8	8, 8	8, 8	8, 8
5	5, 4	7, 8	7, 8	8, 8	7, 8
4	5, 4	7, 8	7, 8	7, 8	8, 8

seller B

$A\backslash B$	8	7	6	5	4
8	8, 8	7, 7	7, 7	7, 7	7, 7
7	7, 8	8, 8	8, 8	8, 8	8, 8
6	7, 8	7, 8	8, 8	8, 8	8, 8
5	7, 8	7, 8	7, 8	8, 8	7, 8
4	7, 8	7, 8	7, 8	7, 8	8, 8

Table 3: Price transition matrices (the first price in each pair corresponds to seller A)

3.4 An Example with unequal equilibrium

In this section we construct an example of an unequal equilibrium, in which prices can converge to a collusive outcome that is not (p_M, p_M). For this example we will use the discrete choice demand function with the following expected payoff from one customer:

$$\pi(p, q) = p \cdot \frac{e^{-bp}}{a + e^{-bp} + e^{-bq}} \tag{30}$$

We choose parameters $a \approx 0.0158$ and $b \approx 0.4760$ so that the monopoly price is $p_M = 8$ and the competitive price is $p_C = 4$. Each seller can choose any prices from the set $\{4, 5, 6, 7, 8\}$. The payoff matrix for these prices with the above parameters is shown in Table 2.

To describe an equilibrium instead of specifying all algorithms that are used (which would be quite complicated) we instead specify the transition matrices, which show which price pair could follow another price pair. Later we can verify that the chosen transition matrices can indeed be sustained in an equilibrium. More specifically, consider the best algorithm that could follow prices (p, q) for seller i (from equation (13)):

$$s^i(p, q) = \arg\max_{s: q \to p} v^i(s) \tag{31}$$

Now denote by $\phi^i(p, q)$ the price pair that would follow (p, q) on the equilibrium sequence starting with $s^i(p, q)$ or more precisely

$$\phi^i(p, q) = \xi^{-i}(s^i(p, q), f^{-i}(s^i(p, q)) \tag{32}$$

If we know the transition matrices $\phi^i(p, q)$ for each seller, we can completely describe how prices will evolve and we can calculate the expected payoffs matrices $v^i(p, q)$, $u^i(p, q)$, $V^i(p, q)$, $U^i(p, q)$. Finally, using Proposition 1 we verify that the specified transitions are feasible and optimal, and, thus, they indeed correspond to an equilibrium. The specific algorithms necessary to sustain such equilibrium can then be recovered from equation (22).

Consider the transition matrices specified in Table 3. We verified that the specified transition matrices can be sustained in an equilibrium for $\beta = 0.9, 0.95, 0.999$ and, also, for any β sufficiently close to 1.[9] For the

[9]We performed the calculations numerically with a Python script. Nevertheless, when comparing two expected payoffs we required

specified transition matrices, depending on the initial condition we end up either with monopoly prices $(8,8)$ or with an unequal outcome $(7,8)$ in which seller A gets a payoff above the monopoly. In order to sustain prices $(7,8)$ in an equilibrium seller A can use the following algorithm:

$$s^A = \{8 \to 7,\ 7 \to 6,\ 6 \to 5,\ 5 \to 4,\ 4 \to 5\} \tag{33}$$

In response to such algorithm, seller B cannot get more than $\pi(8,7)$ immediately and if seller B believes that seller A will keep this algorithm they do not have any incentives to continue with any other price pair. While for A it is optimal to keep this algorithm because it guarantees higher than the monopoly payoff.

4 Precise calculations of expected payoffs

In this extension, we will relax Assumption 2 from Section II.C in the main text, and show that our results generally hold (for a sufficiently small interval between price adjustments Δt) even if we calculate payoffs precisely without ignoring initial convergence and discreteness of price adjustments. First we rewrite the precise model incorporating pre-convergent prices and calculate expected payoffs within this model. Then we study the Markov perfect equilibria of this updated framework.

The main message for this extension can be described in the following informal claim, which is then formalized in the subsequent text:

Claim. *For sufficiently small Δt, every Markov perfect equilibrium in the "precise model" generically corresponds to a Markov perfect equilibrium of the "main model" we study by arbitrarily resolving the dependence on the initial prices.*

4.1 Precise model

Similar to our main model we assume that prices change in the discrete moments $0, 1\cdot\Delta t, 2\cdot\Delta t, \ldots$ with subsequent prices determined by

$$\begin{aligned} p^A_{t+\Delta t} &= s^A(p^B_t) \\ p^B_{t+\Delta t} &= s^B(p^A_t) \end{aligned} \tag{34}$$

where s^A and s^B are the current algorithms of the two sellers. Customers arrive continuously and randomly (also potentially between price adjustments) with Poisson intensity λ. The continuously compounded discount rate is r. Before each price adjustment the seller who was not the last to adjust their algorithm gets the opportunity to adjust their algorithms with probability $(1-e^{-\mu\Delta t})$. When this happens this seller can also choose prices for the current and for the next period.

As before, we *index* all moments in time when one of the sellers adjusts their algorithm by $k = 0, 1, 2, \ldots$, and consider some arbitrary adjustment moment k that happens at time t_k. We express the "precise" expected payoff of the seller adjusting their algorithm, $\tilde{U}_k$, and of the other seller, $\tilde{V}_k$, as follows (note that $\int_0^{\Delta t} \lambda\pi_{k,i}e^{-rt}dt = \frac{\lambda}{r}(1-e^{-r\Delta t})\pi_{k,i}$):

$$\begin{aligned} \tilde{U}_k &= \sum_{i=0}^{\infty} e^{-(r+\mu)i\Delta t}\cdot\left[\tfrac{\lambda}{r}(1-e^{-r\Delta t})\pi_{k,i} + (1-e^{-\mu\Delta t})\tilde{V}_{k+1}\right] \\ \tilde{V}_k &= \sum_{i=0}^{\infty} e^{-(r+\mu)i\Delta t}\cdot\left[\tfrac{\lambda}{r}(1-e^{-r\Delta t})\bar{\pi}_{k,i} + (1-e^{-\mu\Delta t})\tilde{U}_{k+1}\right] \end{aligned} \tag{35}$$

for the difference to be higher than the specified threshold, which was much higher than possible mistakes in the numerical calculations. Thus, we believe the proof is strict. To show that the transition matrices could be sustained as an equilibrium for β sufficiently close to 1 we created a slightly modified version of the script, which compared not the discounted expected payoffs but the sequences of payoffs.

where $\pi_{k,i}$ and $\bar{\pi}_{k,i}$ are the expected payoff from one customer for the corresponding seller in the interval $(t_k + i \cdot \Delta t, t_k + (i+1) \cdot \Delta t)$, in which the prices are constant.

If we denote the average weighted per-period payoffs of each seller as:

$$\pi_k = \frac{\sum_{i=0}^{\infty} e^{-(r+\mu)i\Delta t} \pi_{k,i}}{\sum_{i=0}^{\infty} e^{-(r+\mu)i\Delta t}} \qquad \bar{\pi}_k = \frac{\sum_{i=0}^{\infty} e^{-(r+\mu)i\Delta t} \bar{\pi}_{k,i}}{\sum_{i=0}^{\infty} e^{-(r+\mu)i\Delta t}} \tag{36}$$

and normalize payoffs as $\tilde{u}_k = \frac{\tilde{U}_k}{\lambda/r} \cdot e^{-\mu\Delta t}$ and $\tilde{v}_k = \frac{\tilde{V}_k}{\lambda/r} \cdot e^{-\mu\Delta t}$ then we get the following equations describing the "precise" normalized payoffs:

$$\begin{aligned} \tilde{u}_k &= (1-\tilde{\beta}) \cdot \pi_k + \tilde{\beta} \cdot \tilde{v}_{k+1} \\ \tilde{v}_k &= (1-\tilde{\beta}) \cdot \bar{\pi}_k + \tilde{\beta} \cdot \tilde{u}_{k+1} \end{aligned} \tag{37}$$

Here $\tilde{\beta} = \frac{1-e^{-\mu\Delta t}}{1-e^{-(\mu+r)\Delta t}}$ is the effective discount factor.

Equations (37) are similar to equation set (9) describing the evolution of normalized payoffs in the main model. In addition, $\tilde{\beta} \to \beta = \frac{\mu}{\mu+r}$ as $\Delta t \to 0$. Finally, if $\pi(s_k, s_{k-1})$ and $\bar{\pi}(s_{k-1}, s_k)$ are the limiting expected payoffs from one customer for the two algorithms s_k and s_{k-1}[10], then it follows directly from equations (36) that:

$$\begin{aligned} |\pi_k - \pi(s_k, s_{k-1})| &\le \Delta\pi_{max} \cdot \left(1 - e^{-(r+\mu)\cdot N_k \Delta t}\right) \\ |\bar{\pi}_k - \bar{\pi}(s_{k-1}, s_k)| &\le \Delta\pi_{max} \cdot \left(1 - e^{-(r+\mu)\cdot N_k \Delta t}\right) \end{aligned} \tag{38}$$

where $\Delta\pi_{max} = max_{p,q,p',q'} \left|\pi(p,q) - \pi(p',q')\right|$ and N_k is the length of the convergence period for the current pair of algorithms (which cannot be higher than the number of possible prices squared). Thus, for the same behavior of the sellers, the expected payoffs in the "precise" model (37) would converge to the expected payoffs of the "approximate" limiting model represented by equation (9) in Lamba and Zhuk [2026], as $\Delta t \to 0$.[11]

4.2 Equilibria in the precise model

For sufficiently small Δt the initial price would not have any significant effect on the expected payoff of the seller adjusting their algorithm—by choosing the appropriate initial prices they can always get the continuation sequence with the highest expected payoffs. However, it could be used as a coordination device allowing the seller to select one of the possible best responses with practically identical payoffs. In order not to deal with such situations we assume below that all payoffs in the payoff matrix $\pi(p,q)$ are distinct, that is:

$$\pi(p,q) = \pi(p',q') \quad \Leftrightarrow \quad (p,q) = (p',q') \tag{39}$$

Now consider some arbitrary Markov best response functions for two sellers:

$$\{f^i(s,p),\ p_0^i(s,p),\ p_1^i(s,p)\} \qquad i = A, B \tag{40}$$

Here $f^i(s,p)$ is the algorithm with which seller i would respond and $p_0^i(s,p)$ and $p_1^i(s,p)$ are the initial two prices chosen. Note, the state variable here includes, in addition to the current algorithm of the opponent s, the first price p that the opponent's algorithms will choose. For the pair of responses (40) we can calculate the expected payoffs of the seller adjusting their algorithm $\tilde{u}^i(s,p)$ and of the other seller $\tilde{v}^i(s,p)$ for given values of s and p. Also, denote by $u^i(s,p)$ and $v^i(s,p)$ the limits of these expected payoffs as $\Delta t \to 0$ (for the

[10]To be precise these payoffs also depend on the initial price sequence chosen by the seller adjusting their algorithms. We skip this here for more concise notation.

[11]We again assume sellers cannot respond with an algorithm that would generate a cycle. However, this assumption can be easily relaxed. If we resolve cycles in the main model as equivalent to obtaining the average payoffs over the cycle, then similarly the "precise" payoffs π_k and $\bar{\pi}_k$ would converge to these averages as $\Delta t \to 0$.

same response functions of the sellers (40)):

$$u^i(s,p) = \lim_{\Delta t \to 0} \tilde{u}^i(s,p)$$
$$v^i(s,p) = \lim_{\Delta t \to 0} \tilde{v}^i(s,p) \quad (41)$$

Proposition 2. *Suppose that for any $\varepsilon > 0$ there exists $\Delta t = \Delta t_\varepsilon < \varepsilon$ such that the pair of responses from (40) is an equilibrium in the "precise model" then for any β (except for possibly some finite set of values):*

1. *$u^i(s,p) = u^i(s,p')$ and $v^i(s,p) = v^i(s,p')$ for any prices p and p'*

2. *For arbitrary $p^i(s)$ the strategies*
$$\tilde{f}^i(s) = f^i(s, p^i(s)) \quad (42)$$
form an equilibrium in the "main model" with expected payoffs $u^i(s) = u^i(s,p')$ and $v^i(s) = v^i(s,p')$ for any p'.

Proposition 2 shows that if some pair of best responses (40) survives as an equilibrium for arbitrarily small Δt then we can construct an equilibrium in the main model by resolving the possible dependence on price arbitrarily (equation (42)). Moreover, the expected payoffs in the limiting main model would be equal to the limits of the expected payoffs of the precise model as $\Delta t \to 0$. Since there is a finite number of possible response pairs (40) the next corollary directly follows.

Corollary 1. *For any $\beta = \frac{\mu}{\mu + r}$ (except for possibly some finite set) and for sufficiently small Δt if we take any Markov equilibrium in the "precise" model and resolve dependence on the initial price arbitrarily, then the obtained strategies would be an equilibrium in the "approximate" limiting model.*

Thus, if we observe only collusive equilibria in the main model, then all equilibria in the "precise model" for sufficiently small Δt would also be collusive (except for possibly some non-generic subset of discount factors β).

5 Learning and experimentation

In this section, we discuss how sellers learn the opponent's algorithm. In particular, we argue that within our model, experimentation is relatively inexpensive (its "cost" converges to 0 as $\Delta t \to 0$), and thus the sellers will always choose to discover the opponent's algorithm (at least enough to make a decision). As a result, for Δt small enough, we can assume that the seller is best responding to the opponent's algorithm.

Whenever a seller gets an opportunity to revise their algorithm, they can also spend a short period of time experimenting with the opponent's algorithm (by manually choosing an initial short sequence of prices) before committing to a new algorithm themselves.

Suppose $U^{\text{observable}}$ is the highest payoff the seller currently adjusting their algorithm could get if the opponent's algorithm is directly observable. Now, the considered seller can experiment for the first K periods (for K different price levels) to fully recover the opponent's algorithm. Then, their ultimate payoff is at least:

$$\begin{aligned} U &\geq \textstyle\sum_{j=0}^{K-1} e^{-r \cdot j \Delta t} \cdot \frac{\lambda}{r}(1 - e^{-r\Delta t}) \cdot \pi(\ldots,\ldots) + e^{-r \cdot K \Delta t} \cdot U^{\text{observable}} \\ &\geq U^{\text{observable}} - \frac{\lambda}{r}\left(1 - e^{-r \cdot K \Delta t}\right) \Delta\pi_{\max} \end{aligned} \quad (43)$$

where $\Delta\pi_{\max} = \max_{p,q,p',q'} \left|\pi(p,q) - \pi(p',q')\right|$.[12] The second term in the above expressions converges to 0

[12] In the expression we ignore the possibility of simultaneous algorithm adjustments, which for small Δt is a small probability event, and, thus, would not materially affect the ultimate payoff.

as $\Delta t \to 0$, thus, the actual expected payoffs approximate $U^{\text{observable}}$ for small Δt.[13] If the seller does not experiment and chooses the response immediately, their expected payoff cannot be higher than $U^{\text{observable}}$ and could be sufficiently lower (even for small Δt) because the chosen response might not be optimal for the specific opponent's algorithm.

We should point out two aspects of more general approaches to modeling experimentation. First, algorithms could in principle experiment at arbitrary times to learn more about the opponent's algorithm. At a high level, this would help our conceptual arguments because it would make learning even easier. However, it would make the theoretical analysis intractable by making it harder to separate the pure delegation (to the algorithm) phase from the experimentation phase. Our modeling of the experimentation phase immediately after a revision opportunity makes the analysis tractable while keeping the basic (realistic) forces intact.

Second, for general algorithms with many possible internal states, experimentation could certainly be more costly and it might take more time. It might not even be possible to fully decipher the opponent's algorithm, and you might need to choose your algorithm with some remaining uncertainty. In addition, the opponent's algorithm might have some irreversible state transitions (e.g. some sort of grim-trigger algorithm), thus, it might not be obvious sometimes whether you want to experiment at all.

It would, of course, be interesting and also challenging to study these more general approaches to experimentation in algorithmic pricing. It is beyond the scope of the already detailed analysis here, and we leave it for future research. See Banchio and Mantegazza [2026] and Possnig [2025] for some recent related work on this.

6 Proofs

Proof of Proposition 1

1. First, let us show that all the conditions are necessary for any price pair (p', q') that follows after (p, q). (This implies that $v^i(p, q)$ cannot be above the maximum.)
   - Any algorithm s consistent with (p, q) should map $q \to p$ which gives us the third condition.
   - In addition, when seller $-i$ gets the opportunity to revise their algorithm, they can always keep the pair (p, q) (e.g. by not changing their previous algorithms) and thus $U^{-i}(q', p')$ cannot be below $U^{-i}(q, p)$, which gives the second condition.
   - Finally, since for any s: $u^{-i}(s) \geq \underline{u}^{-i}$ then the first condition should also hold.
2. Now suppose (p', q') maximizes $V^i(p', q')$ and satisfies all the necessary conditions. Then i could use the following algorithm

$$s^i_*(p, q) = \{q \to p, q' \to p', x \to \underline{s}^i(x)\, for\ x \neq q\ and\ x \neq q'\} \tag{44}$$

   to which $-i$ would respond with some s^{-i} such that $\xi(s^i_*(p, q), s^{-i}) = (p', q')$ with the continuation payoff $u^{-i}(s^i_*(p, q)) = U^{-i}(q', p')$. Otherwise:
   - If $\xi(s^i_*(p, q), s^{-i}) = (p, q)$ then $u^{-i}(s^i_*(p, q)) \leq U^{-i}(q, p) \leq U^{-i}(q', p')$
   - If $\xi(s^i_*(p, q), s^{-i}) \notin \{(p, q), (p', q')\}$ then $u^{-i}(s^i_*(p, q)) \leq \underline{u}^{-i} \leq U^{-i}(q', p')$

[13] Since the experimentation phase is not longer than K prices adjustment cycles, it could be even shorter than the period of convergence to a fixed price pair, which we already ignore in our model by Assumption 2.

Proof of Lemma 1

1. Each seller when adjusting their algorithm always has the option to choose the competitive price p_C and keep it forever. Thus $u^i(s_k) \geq \pi(p_C, p_C)$.

2. Consider some $k \geq 1$ suppose $\xi(s_{k-1}, s_k) = (q, p)$. Then

$$u^{-i}(s_{k-1}) = (1-\beta) \cdot \pi(p, q) + \beta \cdot v^{-i}(s_k) \geq \pi(p_C, p_C)$$

3. If $\pi(p, q) \leq \pi(p_C, p_C)$ then $v^{-i}(s_k) \geq \pi(p_C, p_C)$. If $\pi(p, q) > \pi(p_C, p_C)$ then seller $-i$ can choose $s'_k = \{q \to p, x \to x \, for \, x \neq q\}$ and keep it forever, with the payoff of at least $\pi(p_C, p_C)$. Thus, $v^{-i}(s_k) \geq \pi(p_C, p_C)$.

Proof of Lemma 2

1. Note that $v^{-i}(s_k) = (1-\beta) \cdot \bar{\pi}(s_k, s_{k+1}) + \beta \cdot u^{-i}(s_{k+1})$. Then, seller $-i$ can keep the strategy s_k and generate the continuation payoff of at least $\bar{\pi}(s_k, s_{k+1})$, which proves the first statement of the lemma.

2. Note that $u^i(s_k) = (1-\beta) \cdot \pi(s_{k+1}, s_k) + \beta \cdot v^i(s_{k+1})$. Suppose $\xi(s_k, s_{k+1}) = (q, p)$.

3. If $q \neq q^i_*$ then seller i can use the following algorithm[14]

$$s^i = \{s^i_* | q \to p\} \tag{46}$$

the continuation payoff for which is either v^i_* or not lower than $\pi(p, q) = \pi(s_{k+1}, s_k)$.

4. If $q = q^i_*$ then seller i can use the following algorithm

$$s^i = \{s^i_{**} | q \to p\} \tag{47}$$

the continuation payoff for which is either v^i_{**} or not lower than $\pi(p, q) = \pi(s_{k+1}, s_k)$.

Proof of Lemma 3

1. Suppose $u^i(s_k) = v^i(s_{k+1}) = u^i(s_{k+2}) = \cdots = c$. This means that each individual payoff of seller i is equal to c:

$$\pi(s_{k+1}, s_k) = \bar{\pi}(s_{k+1}, s_{k+2}) = \pi(s_{k+3}, s_{k+2}) = \cdots = c \tag{48}$$

2. Now consider the highest individual payoff of seller $-i$ in the sequence:

$$\bar{\pi}(s_k, s_{k+1}), \; \pi(s_{k+2}, s_{k+1}), \; \bar{\pi}(s_{k+2}, s_{k+3}), \; \ldots \tag{49}$$

Then the corresponding seller can keep the strategy, in response to which this highest payoff was generated. This would lead to the same continuation payoff c for seller i and the highest possible constant continuation payoff for $-i$.

[14] Here and below we use the following notation to create a new algorithm s' by slightly adjusting the existing one s:

$$s' = \{s | x \to y\} \quad \Leftrightarrow \quad s'(p) = \begin{cases} s(p) & if \; p \neq x \\ y & if \; p = x \end{cases} \tag{45}$$

3. Assume the contrary. This implies that $\underline{u}^{-i} < \pi(p_M, p_M)$ and $\underline{u}^i < \pi(p_M, p_M)$. This together with Lemma A2 implies that an equilibrium with monopoly prices is possible and after (p_M, p_M) all the subsequent prices are (p_M, p_M) as well.

4. Suppose for some k: $\xi(s_k, s_{k+1}) \neq (p_M, x)$ for some $x < p_M$ then instead of s_{k+1} the corresponding seller could use the following strategy

$$s'_{k+1} = \{s_{k+1} | p_M \rightarrow p_M\} \tag{50}$$

which would lead to the same current payoff and monopoly equilibrium afterwards.

5. If the above k does not exist then for any k for some $x, y < p_M$:

$$\xi(s_k, s_{k+1}) = (p_M, x) \qquad \xi(s_{k+1}, s_{k+2}) = (p_M, y) \tag{51}$$

Then the corresponding seller instead of s_{k+2} can use the following strategy

$$s'_{k+2} = \{s_{k+2} | p_M \rightarrow p_M\} \tag{52}$$

which would generate the current payoff of at least

$$\pi(s'_{k+2}, s_{k+1}) \geq \pi(x, p_M) = \pi(p_M, y) \tag{53}$$

and again the subsequent monopoly equilibrium.

Proof of Lemma 4

We prove the lemma by considering possible cases in terms of the first-best price pairs (p^i_*, q^i_*) from Definition 4. Since we need both price pairs at the same time, to simplify exposition, we use the following notation:

$$\begin{aligned} (p_*, q_*) &= \arg\max V^i(p,q) \qquad s.t.\ U^{-i}(q,p) \geq \underline{u}^{-i} \\ (\dot{q}, \dot{p}) &= \arg\max V^{-i}(q,p) \qquad s.t.\ U^i(p,q) \geq \underline{u}^i \end{aligned} \tag{54}$$

Similarly, we will use (p_{**}, q_{**}) and $(\ddot{q}, \ddot{p})$ as the prices corresponding to the second-best maxima from Definition 5. In general, in this section all prices indexed with letter p (except for $p_C, p_{C+\Delta}, \ldots, p_{M-\Delta}, p_M$) would correspond to seller i and indexed with letter q would correspond to seller $-i$.

For the proof we also need Lemmas A1, A2, A3 stated below and proven after. Note that, if $U^{-i}(q,p) \geq \underline{u}^{-i}$ then seller i can use an algorithm $s = \{q \rightarrow p, x \rightarrow \underline{s}^i(x)\ for\ x \neq q\}$ to generate the expected payoff of at least $\pi(p,q)$, and, thus:

$$\begin{aligned} v^i(p,q) &\geq v^i(s) \geq V^i(p,q) \geq \pi(p,q) \\ u^i(p,q) &\geq U^i(p,q) = (1-\beta)\cdot\pi(p,q) + \beta\cdot v^i(p,q) \geq \pi(p,q) \end{aligned} \tag{55}$$

Similar inequalities hold for seller $-i$ if $U^i(p,q) \geq \underline{u}^i$.

Lemma A 1. *Suppose $\pi(p,q) \geq \pi(p_C,p_C)$ and $\pi(q,p) \geq \pi(p_C,p_C)$ and either $(p,q) \neq (\dot{p}, q_*)$ then either $U^i(p,q) \geq \underline{u}^i$ or $U^{-i}(q,p) \geq \underline{u}^{-i}$ (or both).*

Lemma A 2. *Either $U^i(p_M,p_M) \geq \underline{u}^i$ or $U^{-i}(p_M,p_M) \geq \underline{u}^{-i}$ (or both).*

Lemma A 3. *If $q_* = p_M$ then for any $p \neq p_*$: either $\pi(p_*, q_*) > \pi(p, q_*)$ or $\pi(q_*, p_*) > \pi(q_*, p)$. Moreover, $V^i(p_*, q_*) = \pi(p_*, q_*)$ and $U^{-i}(q_*, p_*) = \pi(q_*, p_*)$.*

Case $q_* \neq p_M$ and $\dot{p} \neq p_M$

From Lemma A2 either $U^i(p_M, p_M) \geq \underline{u}^i$ or $U^{-i}(p_M, p_M) \geq \underline{u}^{-i}$. Suppose for example the latter, then $v^i_{**} \geq V^i(p_M, p_M) \geq \pi(p_M, p_M)$.

Case $q_* = p_M$ and $\dot{p} \neq p_M$

1. We need to consider only $U^i(p_M, p_M) < \underline{u}^i$. Otherwise, $v^{-i}_{**} \geq V^{-i}(p_M, p_M) \geq \pi(p_M, p_M)$.

2. Then, by Lemma A2: $U^{-i}(p_M, p_M) \geq \underline{u}^{-i}$ and, thus, $\underline{u}^i > U^i(p_M, p_M) \geq \pi(p_M, p_M)$.

3. In addition, $p_* \neq p_M$. Otherwise, consider algorithm $s = \{x \to x\}$. For this algorithm $v^{-i}(s) \geq \pi(p_M, p_M)$ and, thus, $\underline{u}^i \leq u^i(s) \leq \pi(p_M, p_M)$.

4. Let us show that $U^{-i}(p_{M-\Delta}, p_{M-\Delta}) \geq \underline{u}^{-i}$

   - If not, then by Lemma A1: $U^i(p_{M-\Delta}, p_{M-\Delta}) \geq \underline{u}^i$ and, thus:

$$\begin{aligned} U^{-i}(p_{M-\Delta}, p_{M-\Delta}) &\geq (1-\beta)\cdot\pi(p_{M-\Delta}, p_{M-\Delta}) + \beta \cdot v^{-i}(p_{M-\Delta}, p_{M-\Delta}) \\ &\geq \pi(p_{M-\Delta}, p_{M-\Delta}) \end{aligned} \tag{56}$$

   - However, by Lemma A3: $\underline{u}^{-i} \leq U^{-i}(q_*, p_*) = \pi(p_M, p_*) < \pi(p_{M-\Delta}, p_*) \leq \pi(p_{M-\Delta}, p_{M-\Delta})$.

5. Finally since $U^{-i}(p_{M-\Delta}, p_{M-\Delta}) \geq \underline{u}^{-i}$:

$$\begin{aligned} v^i_{**} \geq V^i(p_{M-\Delta}, p_{M-\Delta}) &\geq (1-\beta)\cdot\pi(p_{M-\Delta}, p_{M-\Delta}) + \beta \cdot \underline{u}^i \\ &\geq (1-\beta)\cdot\pi(p_{M-\Delta}, p_{M-\Delta}) + \beta \cdot \pi(p_M, p_M) \end{aligned} \tag{57}$$

Case $q_* = p_M$ and $\dot{p} = p_M$

1. In this case the monopoly price pair (p_M, p_M) is an absorbing state.

   - By Lemma A3: $\underline{u}^{-i} \leq U^{-i}(q_*, p_*) \leq \pi(p_M, p_M)$. Similarly, $\underline{u}^i \leq \pi(p_M, p_M)$.
   - By Lemma A2 for one of the sellers (e.g. $-i$): $U^{-i}(p_M, p_M) \geq \underline{u}^{-i}$, which implies: $U^i(p_M, p_M) = (1-\beta)\cdot\pi(p_M, p_M) + \beta \cdot v^i(p_M, p_M) \geq \pi(p_M, p_M) \geq \underline{u}^i$.
   - Finally, $v^{i,-i}(p_M, p_M) \geq \pi(p_M, p_M)$ and $u^{i,-i}(p_M, p_M) \geq U^{i,-i}(p_M, p_M) \geq \pi(p_M, p_M)$, which implies that after (p_M, p_M) all subsequent price pairs are also (p_M, p_M).

2. Thus, $v^{i,-i}(p_M, p_M) = u^{i,-i}(p_M, p_M) = V^{i,-i}(p_M, p_M) = U^{i,-i}(p_M, p_M) = \pi(p_M, p_M)$.

3. Suppose $U^i(p_{M-\Delta}, p_{M-\Delta}) = (1-\beta)\cdot\pi(p_{M-\Delta}, p_{M-\Delta}) + \beta \cdot v^i(p_{M-\Delta}, p_{M-\Delta}) > \pi(p_M, p_M)$

   - Since $v^i(p_{M-\Delta}, p_{M-\Delta}) \leq V^i(p_*, q_*)$ then $p_* \neq p_M$.
   - If $U^{-i}(p_{M-\Delta}, p_{M-\Delta}) < \underline{u}^{-i}$ then by Lemma 1: $U^i(p_{M-\Delta}, p_{M-\Delta}) \geq \underline{u}^i$.
     - As a result, $U^{-i}(p_{M-\Delta}, p_{M-\Delta}) \geq \pi(p_{M-\Delta}, p_{M-\Delta})$.
     - However, by Lemma A3: $\underline{u}^{-i} \leq \pi(q_*, p_*) < \pi(p_{M-\Delta}, p_{M-\Delta})$.

- Finally, if $U^{-i}(p_{M-\Delta}, p_{M-\Delta}) \geq \underline{u}^{-i}$:

$$\begin{aligned} V^i(p_{M-\Delta}, p_{M-\Delta}) &= (1-\beta)\cdot\pi(p_{M-\Delta}, p_{M-\Delta}) + \beta\cdot u^i(p_{M-\Delta}, p_{M-\Delta}) \\ &\geq (1-\beta)\cdot\pi(p_{M-\Delta}, p_{M-\Delta}) + \beta\cdot U^i(p_{M-\Delta}, p_{M-\Delta}) \\ &\geq (1-\beta)\cdot\pi(p_{M-\Delta}, p_{M-\Delta}) + \beta\cdot\pi(p_M, p_M) \end{aligned} \tag{58}$$

4. Thus, we can assume that $U^{i,-i}(p_M, p_M) \geq U^{i,-i}(p_{M-\Delta}, p_{M-\Delta})$, which also implies that $v^{i,-i}(p_{M-\Delta}, p_{M-\Delta}) \geq \pi(p_M, p_M)$.

5. Consider $v^i(p_{M-\Delta}, p_{M-\Delta}) = V^i(p', q')$:

   - If $q' \neq p_M$ then $v^i_{**} \geq V^i(p', q') \geq \pi(p_M, p_M)$.
   - If $q' = p_M$ and $p' < p_M$ then

$$\begin{aligned} V^i(p', q') = (1-\beta)\cdot\pi(p', p_M) + \beta\cdot u^i(p', p_M) \\ U^{-i}(q', p') = (1-\beta)\cdot\pi(p_M, p') + \beta\cdot v^{-i}(p_M, p') \end{aligned} \tag{59}$$

     - Since $U^{-i}(q', p') \geq U^{-i}(p_{M-\Delta}, p_{M-\Delta}) \geq (1-\beta)\cdot\pi(p_{M-\Delta}, p_{M-\Delta}) + \beta\cdot\pi(p_M, p_M)$ and $\pi(p_{M-\Delta}, p_{M-\Delta}) > \pi(p_M, p')$ then $v^{-i}(p_M, p') > \pi(p_M, p_M)$.
     - However, by Lemma 2: $u^i(p', p_M) \geq V^i(p', q') \geq \pi(p_M, p_M)$. Since joint profit is bounded above by $2\pi(p_M, p_M)$, this implies $v^{-i}(p_M, p') \leq \pi(p_M, p_M)$.
   - Finally, if $v^{i,-i}(p_M, p_M) = V^{i,-i}(p_M, p_M) = \pi(p_M, p_M)$ then

$$V^{i,-i}(p_{M-\Delta}, p_{M-\Delta}) = (1-\beta)\cdot\pi(p_{M-\Delta}, p_{M-\Delta}) + \beta\cdot\pi(p_M, p_M) \tag{60}$$

     and by Lemma A1 either $U^i(p_{M-\Delta}, p_{M-\Delta}) \geq \underline{u}^i$ or $U^{-i}(p_{M-\Delta}, p_{M-\Delta}) \geq \underline{u}^{-i}$.

Finally to complete the proof of Lemma 4 and hence Theorem 2, we present the proofs of Lemmas A1, A2 and A3.

Proof of Lemma A1

Suppose $U^{-i}(q, p) < \underline{u}^{-i}$ and $q \neq q_*$. If we consider algorithm $s^i = \{s^i_* | q \to p\}$ for seller i then $v^i(s^i) = v^i_*$, which implies that:

$$U^i(p, q) = (1-\beta)\cdot\pi(p, q) + \beta\cdot v^i(p, q) \geq (1-\beta)\cdot\pi(p_C, p_C) + \beta\cdot v^i_* \geq \underline{u}^i. \tag{61}$$

Proof of Lemma A2

1. By Lemma A1 we only need to consider the case in which $q_* = \dot{p} = p_M$.

2. Assume $U^i(p_M, p_M) < \underline{u}^i$ and $U^{-i}(p_M, p_M) < \underline{u}^{-i}$ then in any equilibrium sequence we should never observe a monopoly price pair (p_M, p_M).

3. Consider some arbitrary equilibrium sequence $s_0, s_1, s_2, \ldots$ Suppose "odd" algorithms correspond, for example, to seller i and, in addition:

$$\begin{aligned} \xi(s_0, s_1) = (x, p_M) \qquad x \neq p_M \\ \xi(s_1, s_2) = (y, z) \qquad z \neq p_M \end{aligned} \tag{62}$$

then the monopoly price pair (p_M, p_M) is possible.

- To prove this first note that:

$$\underline{u}^i \leq u^i(s_0) = (1-\beta)\cdot\pi(p_M, x) + \beta\cdot V^i(y, z) < (1-\beta)\cdot\pi(p_M, p_M) + \beta\cdot v^i_{**} \tag{63}$$

- Now consider the following two unrelated algorithms:

$$\begin{aligned} s^{-i} &= \{\underline{s}^{-i} | p_M \to p_M\} \\ s^i &= \{s^i_{**} | p_M \to p_M\} \end{aligned} \tag{64}$$

- By responding with s^i to s^{-i} the seller i would get at least $(1-\beta)\cdot\pi(p_M, p_M) + \beta\cdot v^i_{**} > \underline{u}^i$ while if the next price pair is not (p_M, p_M) they would get at most $\underline{u}^i$.

4. Consider an equilibrium sequence $s_0, s_1, s_2, \ldots$ starting with $s_0 = s_T = \{x \to x\, for\ any\ x\}$. In this sequence:
   - $\xi(s_0, s_1) = (x_0, x_0) \neq (p_M, p_M)$.
   - $\xi(s_1, s_2) = (x_1, p_M)$: otherwise we can replace s_1 with $s_1' = \{s_1 | p_M \to p_M\}$ and improve payoffs of both sellers in the previous period.
   - $\xi(s_k, s_{k+1}) = (x_k, p_M)$ for $k > 1$: otherwise by the previous argument there should exist a monopoly price pair.
   - But, then this sequence cannot be an equilibrium. The seller with the highest $\pi(x_k, p_M)$ would find optimal to keep the algorithm s_k.

Proof of Lemma A3

1. Consider $V^i(p_*, q_*)$ and the sequence of price pairs that follows (p_*, q_*) in equilibrium:

$$(p_*, q_*) \to (p', q') \to (p'', q'') \to \ldots \tag{65}$$

2. By Lemma 2: $V^i(p_*, q_*) \leq U^i(p', q')$ and $V^{-i}(q', p') \leq U^{-i}(q'', p'')$.
3. In addition for $q_* = p_M$: $U^{-i}(q_*, p_*) \leq V^{-i}(q', p')$
   - $U^i(p_*, q_*) = (1-\beta)\cdot\pi(p_*, q_*) + \beta\cdot v^i_* \geq \underline{u}^i$ since $\underline{u}^i \leq u^i(\{x \to p_c\}) \leq (1-\beta)\cdot\pi(p_C, p_C) + \beta\cdot v^i_*$.
   - As a result, $V^{-i}(q', p') \geq V^{-i}(q_*, p_*) = (1-\beta)\cdot\pi(q_*, p_*) + \beta\cdot u^{-i}(q_*, p_*) \geq (1-\beta)\cdot\pi(q_*, p_*) + \beta\cdot U^{-i}(q_*, p_*)$.
   - Finally, since $U^{-i}(q_*, p_*) = (1-\beta)\cdot\pi(q_*, p_*) + \beta\cdot V^{-i}(q', p')$ then $V^{-i}(q', p') \geq \pi(q_*, p_*)$ and, thus, $V^{-i}(q', p') \geq U^{-i}(q_*, p_*)$.
4. If $U^i(p', q') \leq V^i(p'', q'')$ then since (p_*, q_*) is the first-best price pair for i:

$$\begin{aligned} V^i(p_*, q_*) &= U^i(p', q') = V^i(p'', q'') = \pi(p_*, q_*) = \pi(p', q') \\ U^{-i}(q_*, p_*) &= V^{-i}(q', p') = U^{-i}(q'', p'') = \pi(q_*, p_*) = \pi(q', p') \end{aligned} \tag{66}$$

   - Suppose, there exists $\tilde{p}$ such that $\pi(\tilde{p}, q_*) \geq \pi(p_*, q_*)$ and $\pi(q_*, \tilde{p}) > \pi(q_*, p_*)$.

- Consider $U^{-i}(q_*, \tilde{p}) = (1-\beta)\cdot\pi(q_*, \tilde{p})+\beta\cdot v^{-i}(q_*, \tilde{p})$. Since $-i$ could use the following algorithm:

$$s^{-i} = \{p_* \to q_*, \tilde{p} \to q_*, x \to \underline{s}^{-i}(x)\; for\; x \neq p_*, \tilde{p}\} \tag{67}$$

  then $v^{-i}(q_*, \tilde{p}) \geq v^{-i}(s^{-i}) \geq \pi(q_*, p_*) \geq \underline{u}^{-i}$ and, thus, $U^{-i}(q_*, \tilde{p}) \geq \underline{u}^{-i}$. But, then (p_*, q_*) is not the first best price pair.

5. If $U^i(p', q') > V^i(p'', q'')$ then by Lemma 2 either $(p'', q'') = (p_*, q_*)$ or $(p'', q'') = (p_{**}, q_{**})$ and $q' = q_*$.

6. If $(p'', q'') = (p_*, q_*)$ then

$$\begin{aligned} &U^{-i}(q_*, p_*) = U^{-i}(q', p') = \pi(q_*, p_*) = \pi(q', p') \\ &\pi(p', q') > \pi(p_*, q_*) \end{aligned} \tag{68}$$

   However, then $V^i(p', q') > V^i(p_*, q_*)$ and $U^{-i}(q', p') \geq U^{-i}(q_*, p_*) \geq \underline{u}^{-i}$ and, thus (p_*, q_*) is not the first best price pair.

7. If $(p'', q'') = (p_{**}, q_{**})$ and $q' = q_*$ then:

   - $U^i(p', q') > V^i(p'', q'')$ implies that $\pi(p_*, q_*) < \pi(p', q_*)$.
   - If $\pi(q_*, p') > \pi(q_*, p_*)$ then

$$\begin{aligned} &U^{-i}(q_*, p') = (1-\beta)\cdot\pi(q_*, p') + \beta\cdot v^{-i}(q_*, p') > U^{-i}(q_*, p_*) \geq \underline{u}^{-i} \\ &V^i(p', q_*) = (1-\beta)\cdot\pi(p', q_*) + \beta\cdot u^i(p', q_*) > V^i(p_*, q_*) \end{aligned} \tag{69}$$

     and, thus, again (p_*, q_*) is not the first best price pair.

8. Finally, let us prove that $V^i(p_*, q_*) = \pi(p_*, q_*)$ and $U^{-i}(q_*, p_*) = \pi(q_*, p_*)$.

   - Since $U^i(p_*, q_*) \geq \underline{u}^i$ then $V^{-i}(q', p') \geq V^{-i}(q_*, p_*) \geq \pi(q_*, p_*)$ which in turn implies that $U^{-i}(q_*, p_*) \geq \pi(q_*, p_*)$. In addition, $V^i(p_*, q_*) \geq \pi(p_*, q_*)$.
   - By Assumption 0' part (iv) from Section V in the main text: $(\pi(p_*, q_*), \pi(q_*, p_*))$ lies on the boundary of the convex hull of the payoffs set[15], which means the above inequalities can only hold if all subsequent price pairs are (p_*, q_*).

Proof of Proposition 2

1. Consider one of the sellers choosing the best response for some s and p. For sufficiently small Δt the seller would always respond with an algorithm s' and initial price sequence that would generate the highest possible limiting expected payoff:

$$u(s, p) = \max_{p', s'(.)} (1-\beta)\cdot\pi(p', s(p')) + \beta\cdot v(s', p') \qquad s.t.\; s'(s(p')) = p' \tag{70}$$

   which does not depend on p.

[15] Even if $p_* < \arg\max \tilde{p}\pi(\tilde{p}, p_M)$ then (p_*, q_*) is the point with the highest possible payoff on one seller.

2. If there are several different p' and s' that generate the highest limiting payoff for the seller adjusting their algorithm then with assumption (39) except for possibly finite set of discount factors β these responses also generate exactly the same limiting expected payoff for the other seller:

$$v(s,p) = (1-\beta)\cdot\pi(s(p'),p') + \beta\cdot u(s',p') \tag{71}$$

3. As in the main model, for sufficiently small Δt the seller adjusting the algorithm would also choose the initial prices that maximize the current expected payoff (since any future responses of the opponent generate the same limiting expected payoff irrespective of the future initial prices).

4. Finally, the strategies (42) are an equilibrium in the "approximate" limiting model since any profitable deviation would also be profitable in the "precise" model for sufficiently small Δt.

References

D. Abreu, D. Pearce, and E. Stacchetti. Toward a theory of discounted repeated games with imperfect monitoring. *Econometrica*, 58(5):1041–1063, 1990.

M. Banchio and G. Mantegazza. Artificial intelligence and spontaneous collusion. 2026. forthcoming *American Economic Review*.

R. Lamba and S. Zhuk. Pricing with algorithms. 2026. forthcoming *American Economic Review: Insights*.

C. Possnig. Reinforcement learning and collusion. 2025. University of British Columbia.

B. Salcedo. Pricing algorithms and tacit collusion. University of Western Ontario, 2015.