Northwestern | Kellogg

CMS-EMS
Center for Mathematical Studies in Economics and Management Sciences

Discussion Paper #1613

"Guarantees in Price Experimentation"

Suraj Malladi
Northwestern University

April 2025

Guarantees in Price Experimentation

Suraj Malladi

April 2025

Abstract

A firm faces a persistent demand curve and initially knows little about its shape. It learns about its demand by setting a price each period and observing quantity sold. We characterize the unique, sequentially optimal pricing strategy that maximizes guaranteed profits. This strategy is memoryless and sets prices as if the true demand curve were the flattest conceivable line passing through last period's price and quantity. On path, the firm's prices and profits rise monotonically.

1 Introduction

Firms do not spring into existence knowing the shape of their demand curves. Those with new products or scant data often experiment with various prices over time to learn demand. How exactly should they experiment?

To answer this question, we characterize sequential pricing strategies that maximize the *guaranteed* sum of discounted flow profits in the face of uncertain demand. One interpretation is that an uninformed firm may want to be robust to rich space of possible shapes that its demand curve might take. Alternatively, this objective captures the pricing behavior of an ambiguity averse firm. More broadly, a firm may want to know its guaranteeable profits as a benchmark against which to evaluate other strategies. We show how prices evolve, where they converge, and how much money a firm might leave on the table when it maximizes its profit guarantee.

In the model, a firm faces a persistent, unknown demand curve. Each period, it sets a price and learns demand at that price. The state space is all downward sloping curves satisfying certain slope bounds in a relevant range of prices and a lower bound on sales at some focal price. The firm is forward-looking and solves for a strategy that always maximizes its worst-case continuation profits, no matter the shape of the true demand curve.

The unique, sequentially optimal strategy is to price *as if* the demand curve were linear and passed through the previous period's price and quantity with a slope equal to the known lower bound. Under this strategy, prices rise monotonically over time. If prices ever strictly rise on path, then they remain below any price that would be set by an informed monopolist. Profits also rise, as the sequentially optimal strategy coincides with the myopic strategy on path. Finally, we show a tight bound on how far long-run profits can be from the informed monopoly benchmark, given what the firm knows about consumer price sensitivity. This bound indicates how much money may be left on the table by following the guarantee maximizing strategy. Knowing this bound can help a firm decide between pursuing such a strategy or first learning more about demand through costly consumer surveys or industry analysis.

These results have a positive interpretation. The model rationalizes how firms with market power and limited pricing experience may profitably inflate prices over time despite time-invariant costs and demand. The mechanism is simple. The firm estimates demand conservatively and prices low. It is then likely to make more sales than anticipated, which creates an incentive to raise price. But the price is raised conservatively, based on an overestimate of demand elasticity. Therefore, the firm is likely to be positively surprised by sales again at the new price, and so on.

The results also have a normative interpretation. Management consulting firms encourage clients to price based on industry elasticity ranges. Venture capitalists also prescribe simple pricing policies that rely only on elasticity estimates, e.g., y-combinator suggests startups raise prices to the point where a 5% price hike would drop sales by 20%. We also produce a simple rule of thumb for which firms need only input their elasticity estimates in some relevant range of prices. But our rule enjoys an economic justification: no other policy has a higher profit guarantee.

Finally, the results address two gaps in the rational price experimentation literature in economics, which aims to characterize strategies that maximize a firm's *expected* sum of discounted flow profits. First, little has been said about the price path of a rational firm with a long horizon:

> "Two approaches have appeared to the question of experimentation in the face of a random demand curve... One involves formulating an infinite horizon model in which attention turns to the limiting expectation... A second approach is restricted to two-period models" (Mirman, Samuelson and Urbano, 1993).

The second gap is that, when optimal experimentation does not converge to an informed monopoly price, existing models do not describe how far off and in which direction firms tend to misprice:

> "When adequate learning does not obtain one cannot understand the long-run outcome independently of the priors or the adjustment

> process by which it was reached. An entirely new kind of analysis is called for in these cases: in order to determine the nature of the long-run behaviour of the agent, one needs to characterise the optimal learning strategy. Unfortunately this is possible analytically only in very simple learning problems" (Aghion, Bolton, Harris and Jullien, 1991).

These gaps persist because price experimentation is a multi-arm bandit with correlated arms: trying any price confers information about demand at other nearby prices. It is known that "except in very special cases, such [problems are] unsolvable—–either analytically or numerically" (Francetich and Kreps, 2020). By studying sequentially guarantee-maximizing strategies, we make progress. Such strategies are rational in the sense in which "everyone more or less agrees ... [they consistently maximize] a well-ordered function, such as a utility or profit function" (Becker, 1962).

Outline. Section 2 describes the model. Section 3 characterizes optimal price experimentation. Section 4 considers optimal strategies for more general state spaces. Section 5 concludes with a discussion of the related literature.

2 Model

A *demand curve* is a decreasing function $d : \mathbb{R}_+ \to \mathbb{R}_+$ of price. A firm faces a persistent, unknown demand curve, D. It initially knows only that D is some element of a set Ω. In each period $t = 0, 1, 2 \ldots$, it sets a price p_{t+1}, observes demand at that price, $D(p_{t+1})$, and earns a flow payoff of $p_{t+1} \cdot D(p_{t+1})$.

Assumption **1.** *Let $p_0, q_0 > 0$ and $\underline{\tau} < \overline{\tau} < 0$. Ω is the set of all demand curves d such that $d(p_0) \geq q_0$, and $\frac{d(x)-d(y)}{x-y} \in [\underline{\tau}, \overline{\tau}]$ if $x \neq y$ and $d(x), d(y) > 0$.*

Assumption 1 says the firm only knows that demand is downward sloping, limits to how quickly demand changes with price and a lower bound to sales at some focal price. For example, consider a firm that wants to enter the ready-to-eat cereal industry in an unoccupied niche (e.g., protein-fortified cereal). The firm determines that it can at least make enough sales q_0 to break even at the average price of other branded cereals, p_0. It knows something about how elastic ($\underline{\tau}$) or inelastic ($\bar{\tau}$) demand can get, based on industry level estimates (e.g., Nevo (2001) estimates price elasticities in the cereal industry). The firm knows little else about the shape of its demand curve prior to making a sale.

Histories $h_t \in H_t$ are finite sequences of points discovered on some $d \in \Omega$, $\{(p_i, d(p_i))\}_{i=1}^t$. Paths $\mathbf{h} \in H$ are infinite sequences of points such that any t-truncation of $\mathbf{h}$ is a history in H_t. Let $p_i^{\mathbf{h}}$ be the ith price along path $\mathbf{h} \in H$.

The set of strategies S consists of any function

$$\sigma : \bigcup_{t \in \mathbb{N}} H_t \to \mathbb{R}_+,$$

that maps histories to prices.[1] Starting from a history h_t, a strategy $\sigma \in S$ and a demand curve d together induce a path $\mathbf{h}(\sigma, d, h_t) \in H$, where the t-truncation of $\mathbf{h}(\sigma, d, h_t)$ is h_t.

The firm has a discount rate of $\delta \in (0, 1)$ and constant marginal cost, normalized to zero. Its continuation payoff at any history h_t is the sum of discounted profits,

$$\Pi(\sigma, D, h_t) \equiv \sum_{i=t}^{\infty} \delta^i \cdot p_{i+1}^{\mathbf{h}(\sigma,D,h_t)} \cdot D(p_{i+1}^{\mathbf{h}(\sigma,D,h_t)}).$$

For any pricing strategy $\sigma \in S$, the *guaranteed payoff* to following σ is

$$\min_{d \in \Omega} \Pi(\sigma, d, h_0).$$

[1] Allowing random strategies does not change any results, as they are never optimal. The restriction to deterministic strategies is for ease of exposition only.

Strategy σ^* is *ex-ante optimal* if it maximizes the firm's guaranteed payoff, i.e., σ^* solves

$$\max_{\sigma \in S} \min_{d \in \Omega} \Pi(\sigma, d, h_0). \tag{1}$$

We characterize strategies that are not only ex-ante optimal but maximize the firm's guaranteed continuation payoff at every point in time, given what it learned about D. To that end, let Ω_{h_t} be the set of *consistent* demand curves at h_t, i.e., the demand curves in Ω that pass through the points discovered at h_t. Next, for any strategy σ, say h_t is a *reachable history* if it is the empty history or if there exists $d \in \Omega$ such that h_t is the t-truncation of $\mathbf{h}(\sigma, d, h_0)$. Strategy σ^* is *sequentially optimal* if at every reachable history h_t, it solves

$$\max_{\sigma \in S} \min_{d \in \Omega_{h_t}} \Pi(\sigma, d, h_t). \tag{2}$$

Sequentially optimal strategies are also *ex-ante* optimal, as condition 2 at h_0 is condition 1.

2.1 Discussion

We interpret the firm's environment, information structure and objective.

2.1.1 Absence of strategic considerations

The model applies to settings where learning demand is a primary concern and strategic responses by other firms or consumers are absent or second-order.

The firm may be thought of as a monopolist or as being in a monopolistically competitive industry with a "large enough number [of firms] that they might ignore strategic interactions, a small enough number that they still face downward sloping demand curves" (Stiglitz, 2017).

Similarly, demand can be interpreted as arising from a large population of price-taking consumers with unit demand every period. We can consider goods that are neither durable (so consumers do not delay purchases when price exceeds their value) nor storable (so consumers do not stock up in anticipation of higher prices); or we can suppose consumers are short-lived or myopic.

2.1.2 Alternative assumptions about what the firm knows

Assumption 1 captures a firm that, having not yet entered the market, knows only a lower-bound on quantity demanded at some price. It also says that the firm knows limits to the slope of demand at any price.

We could instead study a firm that already knows a point on its demand curve, and more realistically, knows tighter slope bounds near the focal price and looser bounds farther away. Section 4.2 extends results to such settings.

2.1.3 Sequential Optimality

Sequentially optimal strategies are those that, at every point in time, yield the best guaranteed continuation profit. As the firm learns more about demand, the set of plausible states shrinks, and the worst-case outcome associated with any strategy may change. But a sequentially optimal strategy plans for every contingency at the outset, so the firm would never want to deviate from it.

3 Optimal experimentation

Section 3.1 characterizes the unique sequentially optimal strategy. Section 3.2 shows how prices and profits evolve on path. Section 3.3 and 3.4 discuss where prices and profits converge in the long run, relative to the price set and profit enjoyed by a firm that knows its demand curve. Section 3.5 sketches the arguments driving the results.

3.1 The sequentially optimal strategy

At history $h_t = \{(p_i, q_i)\}_{i=1}^{t}$, let

$$d_{h_t}(p) \equiv q_t + \min\{\underline{\tau}(p - p_t), \overline{\tau}(p - p_t)\},$$

and define the *linear pricing strategy* as

$$\sigma_{lin}(h_t) \equiv \arg\max_{p} p \cdot d_{h_t}(p).$$

This strategy is depicted in Figure 3.1.

There are two senses in which σ_{lin} is simple. First, σ_{lin} is *memoryless*: it depends on any history h_t only through last period's price p_t and quantity q_t.[2] Next, computing $\sigma_{lin}(h_t)$ is elementary: it is the monopoly price if demand were piecewise linear with a single kink at (p_t, q_t), even if such a curve is inconsistent with past data (as in Figure 3.1 (*viii*)).

Proposition 1. *σ_{lin} is a sequentially optimal strategy.*

Section 3.5 gives intuitions for the results, while proofs are in Appendix A.

The sequentially optimal strategy is *essentially unique* if for any two sequentially optimal strategies σ_1, σ_2 and any $d \in \Omega$, both strategies lead down the same path, i.e., $\mathbf{h}(\sigma_1, d, h_0) = \mathbf{h}(\sigma_2, d, h_0)$.

Proposition 2. *The sequentially optimal strategy is essentially unique.*

[2]Formally, let $\lambda(h_t) = (p_t, q_t)$ for any $h_t = \{(p_1, q_1), \ldots, (p_t, q_t)\} \in H_t$. Strategy σ is *memoryless* if it is measurable with respect to λ.

Three Steps of σ_{lin}

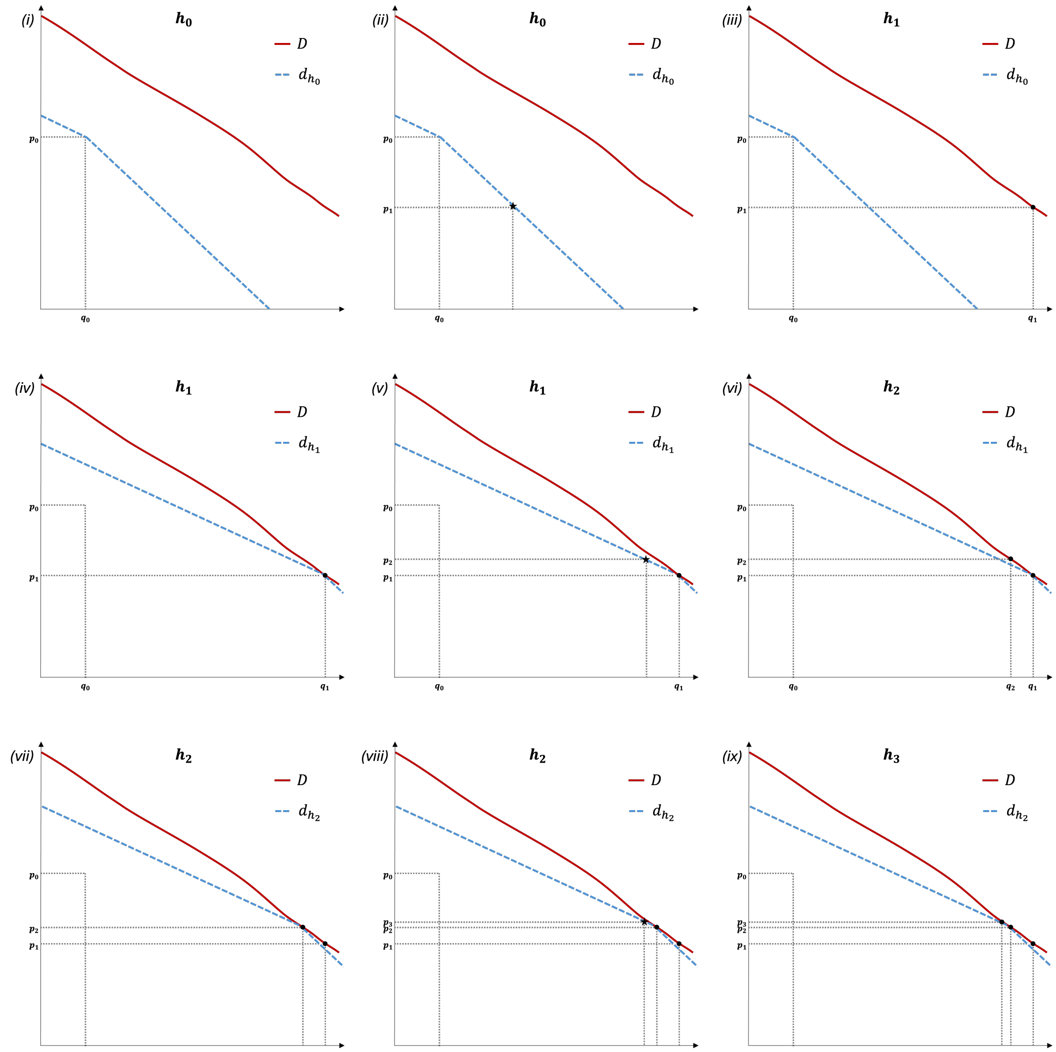


Figure 1: The first row shows the transition from h_0 to h_1 under σ_{lin} when D is the true demand curve. The firm sets the monopoly price p_1 under the assumption that demand is d_{h_0} (top middle) and learns that the actual quantity demanded at p_1 is q_1 (top right). The second row shows the transition from h_1 to h_2, where now the firm sets p_2 under the assumption demand is d_{h_1}. The third row shows the transition from h_1 to h_2. The firm knows d_{h_2} cannot be the true demand curve at h_2, as it does not pass through (p_1, q_1).

Every first course in economics covers monopoly pricing when demand is known and linear. Students often question the value of this exercise when demand curves are often unknown and nonlinear.

Propositions 1 and 2 give one justification. A firm that knows little about its demand curve beyond some bounds to elasticity may want to guarantee itself high profits. While its demand curve may be very irregularly shaped, the unique guarantee maximizing experimentation strategy only involves finding monopoly prices for linear demand curves.

3.2 Price and profit path

We say *prices always rise under strategy* σ if for any $d \in \Omega$, along the path $\mathbf{h} = \mathbf{h}(\sigma, d, h_0)$, $p_i^{\mathbf{h}}$ is increasing in i. Similarly, *profits always rise under strategy* σ if flow profit, $p_i^{\mathbf{h}} \cdot d(p_i^{\mathbf{h}})$, is increasing in i.

Proposition **3.** *Prices and profits always rise under a sequentially optimal strategy.*

In surveys, people frequently cite 'big businesses pursuing profits' or 'greed' as a primary cause for rising prices (Shiller, 1997; Stantcheva, 2024). Based on such responses, Shiller suggests that "most people seem to fail to think of the models that come naturally to economists."

Indeed, it is hard to rationalize rising prices as a consequence of profit maximization, absent any changes to supply and demand. A firm that knows its demand should start and remain at a profit-maximizing price. A firm that is uncertain of its demand might experiment with price hikes *and* cuts, some of which would lower profits as it stumbles to find an optimal price.

However, Proposition 3 shows that maximizing guaranteed profits necessarily involves starting at a lower price and increasing from there. Firms that experiment in this manner enjoy higher profits every step of the way. Finally, these dynamics arise even in the absence of any innovation to a firm's demand or cost, though in that case, the process eventually converges.

3.3 Long-run price

Proposition 3 implies that for any $d \in \Omega$, prices along the path $\mathbf{h}(\sigma, d, h_0)$ generated by a sequentially optimal strategy σ converge to a limit point p_d^∞. The next result describes the relationship between the *long-run price*, p_d^∞, and any *informed monopoly price*, $p_d^* \in \arg\max_p p \cdot d(p)$.

PROPOSITION **4.** *If $\sigma_{lin}(h_0) > p_0$ or $\sigma_{lin}(h_1) > \sigma_{lin}(h_0)$, then $p_D^\infty \leq p_D^*$. Conversely, if $\sigma_{lin}(h_0) \leq p_0$, there is a $d \in \Omega$ such that $p_d^\infty \geq p_d^*$.*

Unless the initial price is set below p_0 and never revised, the firm always sets prices below *any* price it would set if it knew its demand curve.

3.4 Long-run profit

For any $d \in \Omega$, the next result describes the relationship between the *long-run flow profit*, $p_d^\infty \cdot d(p_d^\infty)$, and the *informed monopoly flow profit*, $p_d^* \cdot d(p_d^*)$?

PROPOSITION **5.** *For any $d \in \Omega$,*

$$\frac{p_d^\infty \cdot d(p_d^\infty)}{p_d^* \cdot d(p_d^*)} \geq \frac{4 \cdot \underline{\tau} \cdot \overline{\tau}}{(\underline{\tau} + \overline{\tau})^2},$$

and this bound is tight.

If $\underline{\tau} = \overline{\tau}$, in which case the true demand curve is known with certainty, then the firm achieves the informed monopoly flow profit. In general, when $\underline{\tau} > \overline{\tau}$, the firm is eventually guaranteed only a fraction of the informed monopoly flow profit. The fraction of informed monopoly profits that the firm is guaranteed to eventually capture through experimentation is

$$\inf_{d \in \Omega} \frac{p_d^\infty \cdot d(p_d^\infty)}{p_d^* \cdot d(p_d^*)}.$$

Proposition 5 gives a lower bound on this guarantee in terms of only what it knows about the sensitivity of demand to price. Therefore, the tighter is the firm's estimated range of price elasticity, the larger is the fraction of informed monopoly profits it is guaranteed to capture.

Importantly, this lower bound can be computed prior to any experimentation. Suppose a firm's estimated $\overline{\tau}$ and $\underline{\tau}$ are such that long-run flow profits capture, say, 90% of informed monopoly flow profits, and the firm's discount rate is sufficiently high. Such a firm may find it attractive to follow σ_{lin}, even if its original objective was not to maximize guaranteed profits. On the other hand, suppose the range of estimated slope bounds is wide enough that the firm is only guaranteed 35% of the informed monopoly benchmark. The firm may choose to pursue an alternative pricing strategy or conduct costly consumer surveys and market research to sharpen its estimates of $\underline{\tau}$ and $\overline{\tau}$. In sum, Proposition 5 may be useful for a firm, whether its objective is to maximize its profit guarantees or not.

Next, note that the guaranteed profit ratio in Proposition 5 depends only on the bounds on the slope to the demand curve. It does not depend on the extent to which the firm underestimates demand, $D(p_0) - q_0$, at the focal price, p_0. An implication is that a more accurate estimate of price elasticity gives the firm a better sense of guaranteed profits, while accuracy in estimating quantity demanded does not help in this respect.

3.5 Explanations and remarks

We define a myopic strategy, show that it is sequentially optimal. While it is distinct from the linear pricing strategy and is not memoryless, we show that it coincides with the linear pricing strategy at reachable histories. Propositions 1, 2 and 3 follow from these steps. In the proof of Proposition 4, we show that the linear pricing strategy has a simpler characterization where the kinked curve, d_{h_t}, can be replaced with a linear demand curve at most histories. This fact is useful in proving Proposition 5.

3.5.1 The myopic pricing strategy

At any history h_t, let $\underline{d}_{h_t}$ be the lower envelope of all the $d \in \Omega_{h_t}$. This is the most conservative plausible estimate of the firm's demand at any price, given what it knows at history h_t.

Any price in $\arg\max_p p \cdot \underline{d}_{h_t}(p)$ is a *myopically optimal price* at h_t. A *myopic pricing strategy* sets a myopically optimal price at every history, and $\underline{\sigma}$ denotes one such strategy.

The firm in the model is forward-looking: it maximizes guaranteed payoff, which accounts for the flow of all future profits. On the other hand, the myopic pricing strategy at any point in time maximizes the guaranteed profits in that period only.

Lemma **1.** *Prices and profits always rise under strategy* $\underline{\sigma}$.

The intuition for increasing prices is that if a firm conservatively estimates demand, it starts at a lower price and can only be positively surprised by its actual sales. The unanticipated windfall of inframarginal consumers then makes it more attractive to raise prices. Profits are increasing because the myopic pricing strategy only experiments with a new price if that experiment is guaranteed to yield an improvement in that very period.

LEMMA **2.** *$\underline{\sigma}$ is a sequentially optimal strategy.*

We argue that $\underline{\sigma}$ is sequentially optimal by showing that $(\underline{\sigma}, \underline{d}_{h_t})$ is a saddle point of the continuation payoff, $\Pi(\cdot, \cdot, h_t)$ at every history h_t.

If $D = \underline{d}_{h_t}$, then an optimal strategy is to always choose the corresponding informed monopoly price. This is exactly what $\underline{\sigma}$ does on path.

On the other hand, if the firm follows strategy $\underline{\sigma}$, then by construction, its profit would be weakly higher in period t under any demand curve $D \neq \underline{d}_{h_t}$. Moreover, by lemma 1, its profit in all subsequent periods would be weakly higher as well. Therefore, when the firm plays strategy $\underline{\sigma}$, its profits are minimized when $D = \underline{d}_{h_t}$, proving lemma 2.

3.5.2 Intuition for Propositions 1 and 3

The final step to proving Propositions 1 and 3 is to argue that the linear pricing strategy coincides with the myopic pricing strategy at every reachable history h_t. At most reachable histories, $\underline{d}_{h_t}$ takes on a complex shape. The proof of Proposition 1 shows that at those histories, certain regions of $\underline{d}_{h_t}$ can be 'ironed' to be linear without affecting the monopoly price; see Figure 3.5.2.

3.5.3 Intuition for Proposition 2

The argument for lemma 2 shows that if a strategy maximizes the guaranteed continuation profit at h_t, it chooses a myopically optimal price. Therefore, the sequentially optimal strategy is essentially unique if at every reachable history h_t, there is only one myopically optimal price. This holds at the empty history, because $\underline{d}_{h_0}$ is concave. At any other history, $\underline{d}_{h_t}$ is maximized on $[p_t, \infty)$, by

$\sigma_{lin} = \underline{\sigma}$

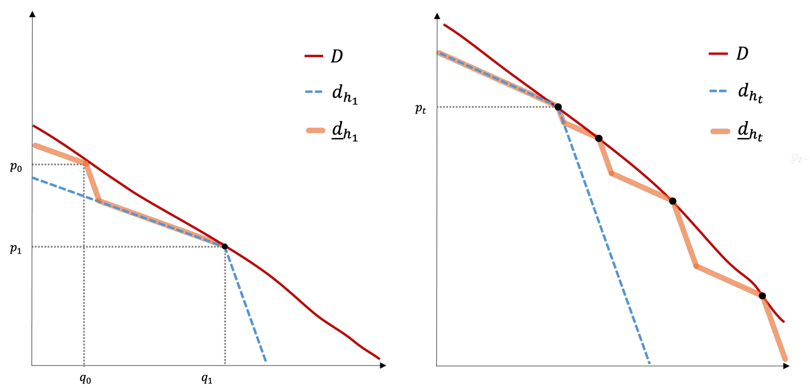


Figure 2: Two kinds of reachable histories under $\underline{\sigma}$ are shown where $d_{h_t} \neq \underline{d}_{h_t}$. The left panel shows a reachable history h_1 at which $p_1 < p_0$. At h_0, price p_1 was chosen over any p where $\underline{d}_{h_1}(p) > d_{h_1}(p)$. Next, $\underline{d}_{h_1}(p) = \underline{d}_{h_0}(p)$ at all such prices. Therefore $\underline{\sigma}(h_1)$ must fall in the region where $\underline{d}_{h_1}$ and d_{h_1} coincide, so $\sigma_{lin}(h_1) = \underline{\sigma}(h_1)$. The right panel shows a reachable history h_t at which $p_t > p_0$. By lemma 1, $\underline{\sigma}(h_t) \geq p_t$. Note that d_{h_t} lies weakly below $\underline{d}_{h_t}$ at $p < p_t$, and both curves coincide at $p \geq p_t$, so $\sigma_{lin}(h_t) = \underline{\sigma}(h_t)$.

$p_D^\infty > p_1$

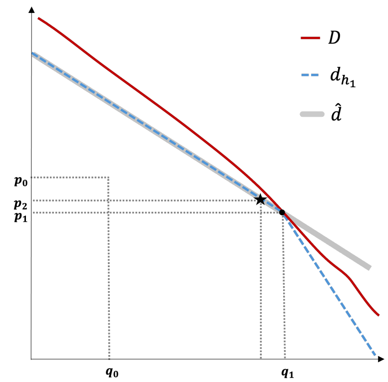


Figure 3: Because p_2 is the optimal price if demand is d_{h_1}, it is also optimal if demand is $\hat{d}$. The true demand curve lies below $\hat{d}$ for prices below p_1, because D is more inelastic everywhere, by Assumption 1. Therefore, $p^\infty > p_1$.

lemma 1. The only reachable histories where $\underline{d}_{h_t}$ is not concave on $[p_t, \infty)$ is when p_1 is sufficiently below p_0, as shown in the left panel of Figure 3.5.2. As argued there, the myopically optimal price is still unique in this case.

3.5.4 Intuition for Proposition 4

Consider the history h_1 in Figure 3.5.4 where $p_2 = \sigma_{lin}(h_1) > p_1$, i.e., p_2 is the optimal price under demand curve d_{h_1}. Clearly, p_2 is also optimal for the linear demand curve $\hat{d}$ that has slope $\underline{\tau}$ and passing through (p_1, q_1). Note, $D(p) < \hat{d}(p)$ if and only if $p < p_1$. So if a price below p_1 is not optimal when demand is $\hat{d}$, it is not optimal when demand is D. Next, at h_2, the linear pricing strategy would set a strictly higher price than p_2. The same logic would show $p_D^* \geq p_2$, and so on. Therefore, $p_D^\infty \leq p_D^*$.

The proof in Appendix A addresses the case where prices never change.

3.5.5 Remark: from kinked to linear demand curves

Figure 3.5.4 shows that if the price strictly increases in the second period, d_{h_t} can thereafter be replaced by the line of slope $\underline{\tau}$ through (p_t, q_t) in the algorithm for computing σ_{lin}. So, after the first period, the sequentially optimal strategy can be described as selecting the monopoly price as if demand were the flattest line passing through the previous period's price and quantity.[3]

3.5.6 Intuition for Proposition 5

From the description in Section 3.5.5, it follows that if prices ever strictly increase on path under σ_{lin}, then p_D^∞ is the monopoly price for the demand curve of slope $\underline{\tau}$ that passes through $(p_D^\infty, D(p_D^\infty))$. By Proposition 4, this long-run price is below any informed monopoly price. This gap is maximized when D is maximally inelastic above the long-run price, i.e., $D(p) = \overline{\tau}(p - p_D^\infty) + D(p_D^\infty)$ for all $p \geq p_D^\infty$. For such a D, the ratio of revenue at p_D^∞ to p_D^* is $\frac{4 \cdot \underline{\tau} \cdot \overline{\tau}}{(\underline{\tau} + \overline{\tau})^2}$.

If D is such that prices never change on path, the proof in Appendix A shows that the fraction of informed monopoly profits captured only goes up.

3.6 Extensions

There are several ways of generalizing the baseline model. For example, it is natural to consider experimentation when demand varies over time or the firm chooses quantities rather than prices (see Appendix B). Section 4 considers a generalization that is closer to our focus. It considers more general state spaces to capture different kinds of knowledge firms may have about the shape of their demand curves prior to experimentation.

[3]This characterization may be reminiscent of Carroll (2015), who finds linear contracts are robustly optimal in a principal-agent problem with moral hazard.

4 Generalizing what the firm may know

In the baseline model, the firm only knows that demand cannot change too quickly and is bounded below at a focal price. Here, we consider firms that that know other things about demand prior to experimentation, e.g., demand at several prices, curvature of demand, the size of the market, etc. Section 4.1 shows by example that the sequentially optimal strategy need not be myopic in general. Section 4.2 characterizes when a myopic pricing strategy is sequentially optimal.

4.1 Myopic strategies are not always optimal

Suppose there are only two possibilities for what the true demand curve might be. In particular, let $d_1(p) = q_0 + \underline{\tau}(p - p_0)$ and $d_2(p) = q_0 + \overline{\tau}(p - p_0)$, and suppose $\Omega = \{d_1, d_2\}$. Let $R_1 = \max_p p \cdot d_1(p)$ and $R_2 = \max_p p \cdot d_2(p)$. Let $p^m \equiv \arg\max_p p \cdot \min\{d_1(p), d_2(p)\}$. See Figure 4.1.

If $p^m > p_0$, then an optimal strategy sets price p_1 in the first period. This is because the informed monopoly price is above p_0 in either state of the world, and in the worst-case, d_1 is the true demand curve. In the second period, the firm learns the true demand curve and sets the informed monopoly price then onward. The case where $p^m < p_0$ is analogous.

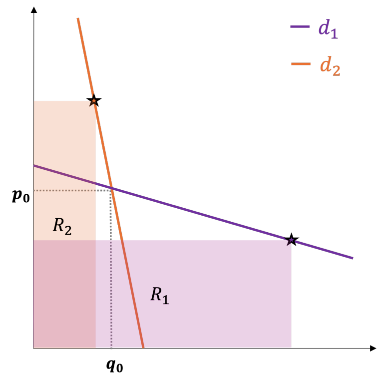
If $p^m = p_0$, the myopic strategy sets a price of p_0 forever, but we argue this is not generally optimal. Suppose $R_1 > R_2$. Let $\bar{p} = \arg\max_p p \cdot d_2(p)$.

Result. *An optimal strategy at h_0 sets a price of*

$$\min\left\{\frac{1}{2} \cdot \left(p_0 + \sqrt{p_0^2 + 4\frac{\delta}{1-\delta}\frac{R_1 - R_2}{\overline{\tau} - \underline{\tau}}}\right),\ \bar{p}\right\}.$$

In particular, a myopic firm ($\delta = 0$) sets an initial price of p_0, while sufficiently patient firms set a price of $\bar{p} > p_0$. More generally, the result implies

Binary state example



that more patient firms experiment farther from the myopically optimal price.

When $R_1 > R_2$, the informed monopoly price is below p_0 if $D = d_1$ and above p_0 if $D = d_2$. Intuitively, the firm hedges by searching closer to the optimal price for the demand curve d_2 that yields the lower informed payoff. If indeed d_2 is the true demand curve, then the firm is better off in the first period for having set a higher price than the myopically optimal p_0. On the other hand, if d_1 happens to be the true demand curve, the firm learns about it and corrects course, whereas the true demand curve is not identified at price p_0. We formalize this argument.

Proof. First, we show that an optimal strategy sets a price $p \in [p_0, \bar{p}]$ at h_0.

Let $p_1 < p_0$. Let σ' and σ'' be the strategies that set a price of p_1 or $\frac{p_0+p_1}{2}$, respectively, at h_0 and subsequently set the informed monopoly price. $\Pi(\sigma', d_2, h_0) < \Pi(\sigma', d_1, h_0)$ and $\Pi(\sigma'', d_2, h_0) < \Pi(\sigma'', d_1, h_0)$. However, $\Pi(\sigma', d_2, h_0) < \Pi(\sigma'', d_2, h_0)$, so

$$\min_{d \in \Omega} \Pi(\sigma', d, h_0) < \min_{d \in \Omega} \Pi(\sigma'', d, h_0).$$

Therefore, no optimal strategy never sets a price strictly below p_0 at h_0. Next, a strategy that sets an initial price above $\bar{p}$ does worse in any state than a strategy that sets a price of $\bar{p}$.

Now, at any initial price $p \in (p_0, \bar{p}]$, the firm immediately learns the true demand curve and sets the informed monopoly price in subsequent periods. Therefore, the profit guarantee to such a strategy is

$$\min\{p \cdot d_1(p) + \frac{\delta}{1-\delta} \cdot R_1, p \cdot d_2(p) + \frac{\delta}{1-\delta} \cdot R_2\}.$$

Maximizing this objective gives the result. □

A rough intuition for why the myopic strategy is not optimal in the binary state example is that experimentation is much more informative. If the firm instead faces a sufficiently rich set of possibilities for demand, it cannot easily discern its shape from the points it discovers. This diminishes incentives to learn about demand, so the firm turns to maximizing guaranteed profits today.

Section 4.2 gives a more complete intuition.

4.2 Environments where myopic strategies are optimal

In the baseline model, there is a worst-case demand curve that is independent of where the firm experiments, namely, the lower envelope of Ω_{h_t}. But, in the example in Section 4.1, the worst-case demand curve depends on the firm's strategy, as the lower envelope of Ω is not itself a consistent demand curve. Using this observation, we characterize state spaces for which myopic strategies are sequentially optimal.

Formally, a set of strictly downward-sloping demand functions $\Omega \subset C(\mathbb{R}_+)$ is a *state space* if

1. Ω is closed under under the supremum norm, and

2. Ω and the inverse demand functions Ω^{-1} are uniformly equicontinuous.[4]

Equicontinuity is the requirement that there are limits to elasticity at any price, whatever the demand curve may be. The assumption that Ω^{-1} is uniformly equicontinuous implies that for any $d \in \Omega$, the set of prices where demand is strictly positive is bounded, so $\max_p p \cdot d(p)$ exists. Assumption 1 satisfies this definition of a state space.

For any state space Ω, we again use $\underline{d}_{h_t}$ to denote the lower envelope of Ω_{h_t} at history h_t and $\underline{\sigma}$ to denote a myopic pricing strategy (see Section 3.5.1).

A state space Ω is *downward-closed* if $\underline{d}_{h_t} \in \Omega_{h_t}$ for any reachable history h_t under $\underline{\sigma}$. A state space Ω is *effectively downward-closed* if for some $p^* \in \arg\max_p p \cdot \underline{d}_{h_t}(p)$, there is a $d \in \Omega_{h_t}$ such that $p^* \in \arg\max p \cdot d(p)$, for any reachable history h_t under $\underline{\sigma}$. A downward-closed Ω is also effectively downward-closed but the converse need not hold.

THEOREM **1.** *If the state space Ω is effectively downward-closed, then some myopic pricing strategy is sequentially optimal, and profits always rise under this strategy. Conversely, if Ω is not effectively downward-closed and the firm is sufficiently patient, i.e., δ is sufficiently close to 1, then the myopic strategy is not sequentially optimal.*

The first part of the statement follows the same steps as in the proofs of Proposition 1 and Proposition 3.

If Ω is not effectively downward-closed, the Arzelà-Ascoli theorem implies that there is a history h_t, reachable under $\underline{\sigma}$, where

$$\max_p \underline{d}_{h_t}(p) < \inf\{\max_p p \cdot d(p) | d \in \Omega_{h_t}\} \equiv m.$$

[4]A set of functions Ω is uniformly equicontinuous if for all $\epsilon > 0$, there exists a $\delta > 0$ such that if $p, p' \in \mathbb{R}_+$ and $|p - p'| < \delta$, then for any $d \in \Omega$, $|d(p) - d(p')| < \epsilon$.

By searching a fine enough grid of prices at h_t, the firm can find a price where flow profit is close to m. If the firm is sufficiently patient, embarking on this grid search at h_t is a profitable deviation from myopic strategy.

4.3 Applications of Theorem 1

Theorem 1 can be used to derive optimal pricing strategies in many cases.

4.3.1 Variants of baseline model

Any Ω that satisfies Assumption 1 is a downward-closed state space. Any set of demand curves that satisfies a variant of Assumption 1 where the bounds on slope can vary with price[5] is also a downward-closed state space.

Assumption 1 considers state spaces where a lower bound is known at a focal price. One can alternatively consider a state space where several points on the demand curve are known *a priori*.

By Theorem 1, myopic strategies are sequentially optimal in such spaces.

4.3.2 Concave demand

Suppose that the firm knew that the demand curve was concave:

Assumption **2.** *Let* $p_0, q_0 > 0$ *and* $\underline{\tau} < \overline{\tau} < 0$. Ω *is the set of all* ***concave*** *demand curves* d *such that* $d(p_0) \geq q_0$, *and* $\frac{d(x)-d(y)}{x-y} \in [\underline{\tau}, \overline{\tau}]$ *if* $x \neq y$ *and* $d(x), d(y) > 0$.

Note that at any history h_t, the set of demand curves Ω_{h_t} is downward closed because the lower envelope of concave functions is also concave. Therefore, by Theorem 1, a myopic strategy is optimal for this state space.

[5] That is, there exist $\underline{\tau}, \overline{\tau} : \mathbb{R}_+ \to \mathbb{R}_+$ such that for all $d \in \Omega$, d' exists *a.e.* and $\underline{\tau}(p) \leq d'(p) \leq \overline{\tau}(p) < 0$.

We solve for the sequentially optimal pricing strategy when the firm knows that demand is concave. Analogues of all the results in the baseline case hold. The strategy is again memoryless, though at some histories, the firm updates the slope of the artificial demand curve used to compute the next price.

At the empty history, h_0, or at any history $h_t = \{(p_i, q_i)\}_{i=1}^t$ where $q_t + \underline{\tau}(p_0 - p_t) > q_0$, let

$$d_{h_t}(p) \equiv q_t + \min\{\underline{\tau}(p - p_t), \overline{\tau}(p - p_t)\},$$

and at any other history $h_t = \{(p_i, q_i)\}_{i=1}^t$, let $\tau = \frac{p_0 - p_t}{q_0 - q_t}$ and

$$d_{h_t}(p) \equiv q_t + \min\{\tau(p - p_t), \overline{\tau}(p - p_t)\}.$$

Let

$$\sigma_{ccv}(h_t) \equiv \arg\max_p p \cdot d_{h_t}(p).$$

Let p_d^∞ denote the long-run price under σ_{cvx} when demand is given by d.

PROPOSITION **6.** *Under assumption 2:*

1. *σ_{ccv} is a sequentially optimal strategy.*

2. *The sequentially optimal strategy is essentially unique.*

3. *Prices and profits always rise under a sequentially optimal strategy.*

4. *If $\sigma_{ccv}(h_0) > p_0$ or $\sigma_{ccv}(h_1) > \sigma_{ccv}(h_0)$, then $p_D^\infty \leq p_D^*$. Conversely, if $\sigma_{ccv}(h_0) \leq p_0$, there is a $d \in \Omega$ such that $p_d^\infty \geq p_d^*$.*

5. *For any $d \in \Omega$,*
$$\frac{p_d^\infty \cdot d(p_d^\infty)}{p_d^* \cdot d(p_d^*)} \geq \frac{4 \cdot \underline{\tau} \cdot \overline{\tau}}{(\underline{\tau} + \overline{\tau})^2},$$
and this bound is tight.

A Ω satisfying Assumption 3 is effectively downward closed

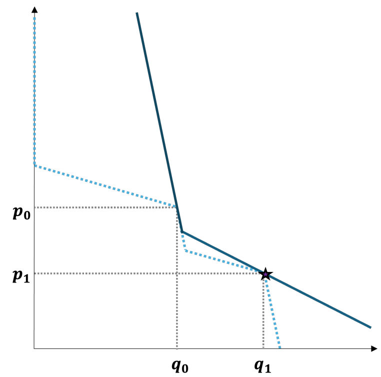


Figure 4: The dashed curve is the lower envelope of some Ω satisfying Assumption 3 at some history. It is not convex, so it is not in Ω. The solid curve satisfies Assumption 3 and is consistent at this history. Revenue on the solid and dashed demand curves are maximized at the same price, p_1.

4.3.3 Convex demand

Suppose instead that the firm knew the demand curve was convex.

Assumption **3.** *Let $p_0, q_0 > 0$ and $\underline{\tau} < \overline{\tau} < 0$. Ω is the set of all* ***convex*** *demand curves d such that $d(p_0) \geq q_0$, and $\frac{d(x)-d(y)}{x-y} \in [\underline{\tau}, \overline{\tau}]$ if $x \neq y$ and $d(x), d(y) > 0$.*

The lower envelope of convex functions is not convex, so any Ω satisfying Assumption 3 is not downward closed. But it is effectively downward closed; see Figure 4.3.3. So by Theorem 1, a myopic strategy is sequentially optimal.

We solve for the sequentially optimal pricing strategy when the firm knows that the demand curve in convex. The strategy has one period memory. The slope of the line between the last two points discovered on the demand curve is a lower bound on the slope of the demand curve at higher prices.

At the empty history, h_0, or at history $h_1 = \{(p_i, q_i)\}$, or at a history

$h_t = \{(p_i, q_i)\}_{i=1}^t$ where $p_t = p_{t-1}$, let

$$d_{h_t}(p) \equiv q_t + \min\{\underline{\tau}(p - p_t), \overline{\tau}(p - p_t)\},$$

and at any other history $h_t = \{(p_i, q_i)\}_{i=1}^t$, let $\tau = \frac{p_t - p_{t-1}}{q_t - q_{t-1}}$ and

$$d_{h_t}(p) \equiv q_t + \tau(p - p_t).$$

Let

$$\sigma_{cvx}(h_t) \equiv \arg\max_p p \cdot d_{h_t}(p).$$

Let p_d^∞ denote the long-run price under σ_{cvx} when demand is given by d.

PROPOSITION **7.** *Under assumption 3:*

1. σ_{cvx} *is a sequentially optimal strategy.*

2. *The sequentially optimal strategy is essentially unique.*

3. *Prices and profits always rise under a sequentially optimal strategy.*

4. *If* $\sigma_{cvx}(h_0) > p_0$ *or* $\sigma_{cvx}(h_1) > \sigma_{cvx}(h_0)$*, then* $p_D^\infty \leq p_D^*$*. Conversely, if* $\sigma_{cvx}(h_0) \leq p_0$*, there is a* $d \in \Omega$ *such that* $p_d^\infty \geq p_d^*$*.*

5. *For any* $d \in \Omega$*,*
$$\frac{p_d^\infty \cdot d(p_d^\infty)}{p_d^* \cdot d(p_d^*)} \geq \frac{4 \cdot \underline{\tau} \cdot \overline{\tau}}{(\underline{\tau} + \overline{\tau})^2},$$
and this bound is tight.

5 Related literature

This paper contributes to a literature on rational price experimentation with uncertain demand. The standard approach considers a firm that maximizes expected discounted flow profits for some distribution over demand curves.

Here, we consider a firm that maximizes guaranteed discounted flow profits instead. Such a firm's prices and profits rise over time.

Sequential price experimentation is a bandit problem with correlated arms, so deriving optimal strategies for a long-lived firm is typically intractable. We describe how others tackled this problem and how our approach compares.

One approach derives conditions on state spaces and beliefs under which optimal experimentation converges to the price or profit of an informed monopolist (Rothschild, 1974; McLennan, 1984; Easley and Kiefer, 1988; Aghion et al., 1991). Easley and Kiefer (1988) consider noisy learning, i.e., a stochastic demand curve. Aghion et al. (1991) focus on deterministic learning, and we follow their approach. These papers show where optimal strategies converge without constructing them explicitly. Accordingly, their results do not have normative implications for experimentation. One may also question the positive implications about long-run learning: why should a firm be able to solve for optimal strategies when the analyst cannot?

Another approach studies price experimentation with shorter planning horizons; see Mirman et al. (1993) and references therein for references on two period models. Bergemann and Schlag (2011) characterize how an uninformed monopolist prices to minimize maximum regret in a static model. Handel and Misra (2015) study a two period version of this problem. They also ask how new firms set prices and focus on a notion of sequential optimality.[6] By contrast, our goal is to characterize both the path and limit of

[6]Other papers also study sequential optimality in settings with minmax preferences. Hanany et al. (2020) apply this refinement to games with ambiguity averse agents. Libgober and Mu (2021) and Li et al. (2022) study pricing for a durable good with robustness to buyer learning. Malladi (2022) and Banchio and Malladi (2024) study search with learning. We take this approach to a bandit problem with uncertain flow payoffs and no stopping.

optimal price experimentation for a long-lived and forward-looking firm.

A different approach partially characterizes optimal policies assuming a binary state space (Rustichini and Wolinsky, 1995; Keller and Rady, 1999) or with strong parametric assumptions on demand (e.g., that demand is linear or log-linear or isoelastic). By contrast, we try to make credible assumptions about the richness of the space of demand curves a firm may face and the extent of its uncertainty.[7]

Finally, there is a literature with roots in operations research that develops pricing algorithms that are practically useful or perform well in the limit, even if they do not optimally account for intertemporal incentives in experimentation; see the review by Den Boer (2015). As such, this literature focuses on approximate solutions to complex dynamic programs (e.g., Lobo and Boyd (2003)), restricts attention to simple pricing rules (e.g., Cohen et al. (2021); Cho and Libgober (2022)), or consider criteria like asymptotic regret that place no restriction on short-run behavior (e.g., Besbes and Zeevi (2009)). By contrast, we study a rational firm (i.e., profit maximizing, forward looking, and sequentially optimal) to understand price experimentation for new products.

[7] Kang and Vasserman (2022) make similar assumptions to model a policy-maker's uncertainty about demand when considering interventions like taxes or subsidies.

References

Aghion, Philippe, Patrick Bolton, Christopher Harris, and Bruno Jullien, "Optimal learning by experimentation," *The review of economic studies*, 1991, *58* (4), 621–654.

Banchio, Martino and Suraj Malladi, "Search and Rediscovery," in "Proceedings of the 25th ACM Conference on Economics and Computation" 2024, pp. 1194–1194.

Becker, Gary S, "Irrational behavior and economic theory," *Journal of political economy*, 1962, *70* (1), 1–13.

Bergemann, Dirk and Karl Schlag, "Robust monopoly pricing," *Journal of Economic Theory*, 2011, *146* (6), 2527–2543.

Besbes, Omar and Assaf Zeevi, "Dynamic pricing without knowing the demand function: Risk bounds and near-optimal algorithms," *Operations research*, 2009, *57* (6), 1407–1420.

Boer, Arnoud V Den, "Dynamic pricing and learning: historical origins, current research, and new directions," *Surveys in operations research and management science*, 2015, *20* (1), 1–18.

Carroll, Gabriel, "Robustness and linear contracts," *American Economic Review*, 2015, *105* (2), 536–563.

Cho, In-Koo and Jonathan Libgober, "Learning Underspecified Models," *arXiv preprint arXiv:2207.10140*, 2022.

Cohen, Maxime C, Georgia Perakis, and Robert S Pindyck, "A simple rule for pricing with limited knowledge of demand," *Management Science*, 2021, *67* (3), 1608–1621.

Easley, David and Nicholas M Kiefer, "Controlling a stochastic process with unknown parameters," *Econometrica: Journal of the Econometric Society*, 1988, pp. 1045–1064.

Francetich, Alejandro and David Kreps, "Choosing a good toolkit, II: Bayes-rule based heuristics," *Journal of Economic Dynamics and Control*, 2020, *111*, 103814.

Hanany, Eran, Peter Klibanoff, and Sujoy Mukerji, "Incomplete information games with ambiguity averse players," *American Economic Journal: Microeconomics*, 2020, *12* (2), 135–187.

Handel, Benjamin R and Kanishka Misra, "Robust new product pricing," *Marketing Science*, 2015, *34* (6), 864–881.

Kang, Zi Yang and Shoshana Vasserman, "Robust bounds for welfare analysis," Technical Report, National Bureau of Economic Research 2022.

Keller, Godfrey and Sven Rady, "Optimal experimentation in a changing environment," *The review of economic studies*, 1999, *66* (3), 475–507.

Li, Zihao, Jonathan Libgober, and Xiaosheng Mu, "Sequentially Optimal Pricing under Informational Robustness," *arXiv preprint arXiv:2202.04616*, 2022.

Libgober, Jonathan and Xiaosheng Mu, "Informational robustness in intertemporal pricing," *The Review of Economic Studies*, 2021, *88* (3), 1224–1252.

Lobo, Miguel Sousa and Stephen Boyd, "Pricing and learning with uncertain demand," in "INFORMS revenue management conference" Citeseer 2003, pp. 63–64.

Malladi, Suraj, "Searching in the dark and learning where to look," *Available at SSRN 4084113*, 2022.

McLennan, Andrew, "Price dispersion and incomplete learning in the long run," *Journal of Economic dynamics and control*, 1984, *7* (3), 331–347.

Mirman, Leonard J, Larry Samuelson, and Amparo Urbano, "Monopoly experimentation," *International Economic Review*, 1993, pp. 549–563.

Nevo, Aviv, "Measuring market power in the ready-to-eat cereal industry," *Econometrica*, 2001, *69* (2), 307–342.

Rothschild, Michael, "A two-armed bandit theory of market pricing," *Journal of Economic Theory*, 1974, *9* (2), 185–202.

Rustichini, Aldo and Asher Wolinsky, "Learning about variable demand in the long run," *Journal of economic Dynamics and Control*, 1995, *19* (5-7), 1283–1292.

Shiller, Robert J, "Why do people dislike inflation?," in "Reducing inflation: Motivation and strategy," University of Chicago Press, 1997, pp. 13–70.

Stantcheva, Stefanie, "Why do we dislike inflation?," *Brookings Papers on Economic Activity*, 2024, *2004* (1), 1–46.

Stiglitz, Joseph E, "Monopolistic competition, the Dixit–Stiglitz model, and economic analysis," *Research in Economics*, 2017, *71* (4), 798–802.

A Proofs

Proof of Lemma 1. Let $d \in \Omega$, and let $h_t = \{(p_1, d(p_1)), \ldots (p_t, d(p_t))\}$ be the t-truncation of $\mathbf{h}(\underline{\sigma}, d, h_0)$, for $t \geq 1$. Similarly, let h_{t-1} be the $t-1$ truncation. Let $\underline{\sigma}(h_t) \equiv p_{t+1}$. Finally, let $\epsilon = d(p_t) - \underline{d}_{h_t}(p_t)$.

By construction, $\epsilon \geq 0$ and $\underline{d}_{h_{t-1}} + \epsilon \geq \underline{d}_{h_t}$, so for $p < p_t$,

$$p_t \cdot \underline{d}_{h_t}(p_t) = p_t \cdot (\underline{d}_{h_{t-1}}(p_t) + \epsilon) > p \cdot (\underline{d}_{h_{t-1}}(p) + \epsilon) \geq p \cdot \underline{d}_{h_t}(p),$$

where the first inequality follows from the definition of p_t as the lowest price in $\arg\max_p p \cdot \underline{d}_{h_{t-1}}$. Because the flow profit for demand $\underline{d}_{h_t}$ is not maximized at a price below p_t, we conclude that $\underline{\sigma}(h_t) = p_{t+1} \geq p_t = \underline{\sigma}(h_{t-1})$. Therefore, prices always rise under $\underline{\sigma}$.

Next, note that

$$p_{t+1} \cdot d(p_{t+1}) \geq p_{t+1} \cdot \underline{d}_{h_t}(p_{t+1}) \geq p_t \cdot \underline{d}_{h_t}(p_t) = p_t \cdot d(p_t),$$

where the first inequality is by definition of $\underline{d}_{h_t}$, the second inequality is by definition of $p_{t+1} = \underline{\sigma}(h_t)$. Therefore, profits always rise under $\underline{\sigma}$.

□

LEMMA **3.** *For any history h_t and strategy σ, $\Pi(\sigma, \underline{d}_{h_t}, h_t) \leq \Pi(\underline{\sigma}, \underline{d}_{h_t}, h_t)$.*

Proof of Lemma 3. If $D = \underline{d}_{h_t}$, then $\underline{\sigma}$ sets the same myopically optimal price p' in every subsequent period, for a per period flow profit $p'\underline{d}_{h_t}(p')$. Because p' is also the informed monopoly price, no other strategy generates a higher continuation payoff in this state of the world. □

LEMMA **4.** *For any history h_t and $d \in \Omega_{h_t}$, $\Pi(\underline{\sigma}, \underline{d}_{h_t}) \leq \Pi(\underline{\sigma}, d)$.*

Proof of Lemma 4. Let $p' \equiv \underline{\sigma}(h_t)$. If $d(p') = \underline{d}_{h_t}(p')$, then $\Pi(\underline{\sigma}, d) = \Pi(\underline{\sigma}, \underline{d}_{h_t})$ because prices and flow profits remain constant. Otherwise, flow profits rise monotonically over time by lemma 1.

□

Proof of Lemma 2. Take any history h_t.

By lemma 3, for any strategy $\sigma \in S$, $\Pi(\underline{\sigma}, \underline{d}_{h_t}, h_t) \geq \Pi(\sigma, \underline{d}_{h_t}, h_t)$, i.e.,

$$\Pi(\underline{\sigma}, \underline{d}_{h_t}, h_t) \geq \max_{\sigma \in S} \Pi(\sigma, \underline{d}_{h_t}, h_t) \geq \max_{\sigma \in S} \min_{d \in \Omega_{h_t}} \Pi(\sigma, d, h_t).$$

By lemma 4, for any demand $d \in \Omega_{h_t}$, $\Pi(\underline{\sigma}, \underline{d}_{h_t}, h_t) \leq \Pi(\underline{\sigma}, d, h_t)$, i.e.,

$$\Pi(\underline{\sigma}, \underline{d}_{h_t}, h_t) \leq \min_{d \in \Omega_{h_t}} \Pi(\underline{\sigma}, d, h_t) \leq \max_{\sigma \in S} \min_{d \in \Omega_{h_t}} \Pi(\sigma, d, h_t).$$

Therefore, $\Pi(\underline{\sigma}, \underline{d}_{h_t}, h_t) = \max_{\sigma \in S} \min_{d \in \Omega_{h_t}} \Pi(\sigma, d, h_t)$, so $\underline{\sigma}$ is optimal. □

Proof of Proposition 1. By lemma 2, it suffices to show that on all the reachable histories h_t of $\underline{\sigma}$, $\sigma_{lin}(h_t) = \underline{\sigma}(h_t)$. Note that this also implies that σ_{lin} and $\underline{\sigma}$ share the same reachable histories.

Case 1: $h_t = h_0$. At the empty history,

$$d_{h_0}(p) = q_0 + \min\{\underline{\tau}(p - p_0), \overline{\tau}(p - p_0)\} = \underline{d}_{h_0}(p).$$

Therefore, $\sigma_{lin}(h_0) = \underline{\sigma}(h_0)$.

Case 2: h_t, where $q_t + \underline{\tau}(p_0 - p_t) \geq q_0$. By construction, $\underline{d}_{h_t} \geq d_{h_t}$ for $p < p_t$, and $\underline{d}_{h_t} = d_{h_t}$ for $p \geq p_t$. Next, by lemma 1, $\underline{\sigma}(h_t) \geq p_t$. Therefore, $\sigma_{lin}(h_t) = \underline{\sigma}(h_t)$.

Case 3: h_t, where $q_t + \underline{\tau}(p_0 - p_t) < q_0$. Again, $\underline{d}_{h_t} \geq d_{h_t}$ for $p < p_t$, and $\underline{\sigma}(h_t) \geq p_t$. But for $p \geq p_t$, note that $\underline{d}_{h_t} = \max\{d_{h_t}, \underline{d}_{h_0}\}$.

However, at any $p \geq p_t$ for which $\underline{d}_{h_t}(p) = \underline{d}_{h_0}(p)$,

$$p \cdot d_{h_t}(p) \leq p \cdot \underline{d}_{h_0}(p) \leq p_1 \underline{d}_{h_0}(p_1) < p_t \cdot \underline{d}_{h_{t-1}}(p_t) \leq p_t \cdot \underline{d}_{h_t}(p_t),$$

where the second inequality is by definition of $\underline{\sigma}$, the third inequality is lemma 1, and the fourth is because $\underline{d}_{h_i}$ is uniformly non-decreasing in i. So, at $p_{t+1} = \underline{\sigma}(h_t)$, we have $\underline{d}_{h_t}(p_{t+1}) = d_{h_t}(p_{t+1})$. Therefore, $\sigma_{lin}(h_t) = \underline{\sigma}(h_t)$. □

Proof of Proposition 2. The proof of lemma 3 shows that if at any reachable history h_t for $\underline{\sigma}$, if σ does not solve $\arg\max_p p \cdot \underline{d}_{h_t}(p)$, then σ is not optimal. Therefore, it suffices to show that, at any reachable history h_t for $\underline{\sigma}$, the set $\arg\max_p p \cdot \underline{d}_{h_t}(p)$ is a singleton.

At the empty history, $\underline{d}_{h_0} = d_{h_0}$ is convex, so $\max_p p \cdot \underline{d}_{h_0}(p)$ has a unique solution.

Suppose that at reachable history h_t, price p_t is the unique solution to $\max_p p \cdot \underline{d}_{h_t}(p)$. Let h_{t+1} be a history where price p_t was set at h_t and the corresponding demand was realized.

By lemma 1, if $p \in \arg\max_p p \cdot \underline{d}_{h_{t+1}}(p)$, then $p \geq p_t$, and the proof of lemma 1 shows that the solution is unique on this range of prices. $\square$

Proof of Proposition 3. This follows immediately from lemma 1, lemma 2 and Proposition 2. $\square$

Proof of Proposition 4. Let $\sigma_{lin}(h_0) \equiv p_1 > p_0$. This means for any $p \geq 0$,

$$p_1 \cdot (q_0 + \underline{\tau}(p_1 - p_0)) \geq p \cdot (q_0 + \underline{\tau}(p - p_0)). \tag{3}$$

Next, note that

$$D(p_1) - \underline{\tau}p_1 - q_0 + \underline{\tau}p_0 \geq q_0 + \underline{\tau}(p_1 - p_0) - \underline{\tau}p_1 - q_0 + \underline{\tau}p_0 = 0.$$

Therefore, for $p < p_1$,

$$p_1 \cdot (D(p_1) - \underline{\tau}p_1 - q_0 + \underline{\tau}p_0) \geq p \cdot (D(p_1) - \underline{\tau}p_1 - q_0 + \underline{\tau}p_0). \tag{4}$$

Adding inequalities 3 and 4, we conclude that for any $p < p_1$

$$p_1 \cdot D(p_1) = p_1 \cdot (D(p_1) + \underline{\tau}(p_1 - p_1)) > p \cdot (D(p_1) + \underline{\tau}(p - p_1)) \geq p \cdot D(p).$$

In other words, $p_1 \leq p_D^*$.

Now suppose $\sigma_{lin}(h_t) \equiv p_t \leq p_D^*$, where h_t is the t-truncation of $\mathbf{h}(\sigma, D, h_0)$ for $t \geq 1$. If $\sigma_{lin}(h_{t+1}) \equiv p_{t+1} = p_t$, then $p_{t+1} \leq p_D^*$ by assumption. If $p_{t+1} > p_t$, the proof follows by analogous steps to the base case, so $p_D^\infty \leq p_D^*$. The case where $p_1 < p_0$ but $p_2 > p_1$ also follows the same steps as in the base case.

On the other hand, if $p_1 \leq p_0$, let $d(p) = d_{h_0}(p)$ for $p \geq p_1$ and let $d(p) = d_{h_0}(p_1) + \underline{\tau}(p - p_1)$ for $p < p_1$. If $D = d$, then $p_D^\infty = p_1$ while $p_D^* \leq p_1$.

□

Proof of Proposition 5. For any $d \in \Omega$, let

$$R_d \equiv \frac{p_d^\infty \cdot d(p_d^\infty)}{p_d^* \cdot d(p_d^*)}.$$

Let $p_1 \equiv \sigma_{lin}(h_0)$. There are two cases to consider.

Case 1: D is such that $p_D^\infty > p_1$ or $p_1 > p_0$.

In this case, $p_D^\infty = \arg\max_p p \cdot (D(p_D^\infty) + \underline{\tau}(p - p_D^\infty))$. That is, for some $b > 0$, $p_D^\infty = \arg\max_p p \cdot (\underline{\tau} \cdot p + b)$. In other words, $p_D^\infty = \frac{-b}{2\underline{\tau}}$ for some $b > 0$, and $(\frac{-b}{2\underline{\tau}}, \frac{b}{2})$ is a point on D. Let

$$d(p) = \overline{\tau}(p + \frac{b}{2\underline{\tau}}) + \frac{b}{2}.$$

Note that $D(p) \leq d(p)$ for all $p \geq p_D^\infty$, and $p_D^* \geq p_D^\infty$ by Proposition 4. Therefore,

$$R_D \geq \frac{p_D^\infty \cdot D(p_D^\infty)}{\max_p p \cdot d(p)}.$$

Now, $\arg\max_p p \cdot d(p) = \frac{-b}{4\overline{\tau}}(1 + \frac{\overline{\tau}}{\underline{\tau}})$. Therefore,

$$\begin{aligned}
\max_p p \cdot d(p) &= \frac{-b}{4\overline{\tau}}(1 + \frac{\overline{\tau}}{\underline{\tau}}) \cdot (\overline{\tau}(\frac{-b}{4\overline{\tau}}(1 + \frac{\overline{\tau}}{\underline{\tau}}) + \frac{b}{2\underline{\tau}}) + \frac{b}{2}) \\
&= \frac{-b^2}{4\overline{\tau}}(1 + \frac{\overline{\tau}}{\underline{\tau}}) \cdot (\frac{-1}{4}(1 + \frac{\overline{\tau}}{\underline{\tau}}) + \frac{\overline{\tau}}{2\underline{\tau}}) + \frac{1}{2}) \\
&= \frac{-b^2}{16\overline{\tau}}(1 + \frac{\overline{\tau}}{\underline{\tau}})^2
\end{aligned}$$

Therefore,

$$R_D \geq \frac{p_D^\infty \cdot D(p_D^\infty)}{\max_p p \cdot d(p)} = \frac{4 \cdot \underline{\tau} \cdot \overline{\tau}}{(\underline{\tau} + \overline{\tau})^2}.$$

Case 2: D is such that $p_D^\infty = p_1$ and $p_1 \leq p_0$. We break this case into two sub-cases. Let $d_t(p) = q_0 + \overline{\tau}(p - p_0)$. Next, let $d_b(p) = d_t(p_1) + \underline{\tau}(p - p_1)$. Finally, let $d = \max\{d_t, d_b\}$.

Sub-case 1: $D(p_1) = d_t(p_1)$. In this case, $p_1 = \arg\max_p p \cdot (D(p_1) + \overline{\tau}(p - p_1))$. By arguments symmetric to those in Case 1, the same conclusion follows.

Sub-case 2: $D(p_1) > d_t(p_1)$. In this case, let $\delta > 0$ be such $\arg\max_p p \cdot (d_b(p) + \delta) = p_1$. Let $x \equiv D(p_1) - d_t(p_1)$, and let

$$R(x) \equiv \frac{p_1 \cdot (d(p_1) + x)}{\max_p p \cdot (d(p) + x)}.$$

It cannot be the case that $x > \delta$, as $p_D^\infty > p_1$ in that case.

If $x = \delta$, then $p_D^\infty = \arg\max_p p \cdot (D(p_D^\infty) + \overline{\tau}(p - p_D^\infty))$, so just as in Case 1 or Sub-case 1, $R_D \geq R(\delta) = \frac{4 \cdot \underline{\tau} \cdot \overline{\tau}}{(\underline{\tau} + \overline{\tau})^2}$.

Finally, suppose that $x \in (0, \delta)$.There is an $x' > 0$ such that,

$$\arg\max_p p \cdot (d(p) + x) = \begin{cases} \arg\max_p p \cdot (d_b(p) + x) \text{ if } x \leq x' \\ \arg\max_p p \cdot (d_t(p) + x) \text{ if } x \geq x' \end{cases}$$

If $x \leq x'$, let $p_b^*(x) = \arg\max_p p \cdot (d_b(p) + x)$. Note that if $x \leq x'$, then $p_b^*(x) \leq p_1$, so

$$\begin{aligned} \max_p p \cdot (d_b(p) + x) - \max_p p \cdot d_b(p) &< p_b^*(x) \cdot (d_b(p_b^*(x)) + x) - p_b^*(x) \cdot d_b(p_b^*(x)) \\ &= p_b^*(x) * x \\ &\leq p_1 \cdot (d_b(p_1) + x) - p_1 \cdot d_b(p_1). \end{aligned}$$

Therefore, if $x \leq x'$, then $R(x) \geq R(0) = \frac{4 \cdot \underline{\tau} \cdot \overline{\tau}}{(\underline{\tau}+\overline{\tau})^2}$, where the equality is from Sub-case 1.

Next for $x \geq x'$,

$$\max_p p \cdot (d(p) + x) = \max_p p \cdot (d_t(p) + x) = \frac{-(q_0 - \overline{\tau} p_0 + x)^2}{4\overline{\tau}},$$

so

$$\begin{aligned} R(x) &= (4\overline{\tau} p_1) \cdot \frac{(q_0 + \overline{\tau}(p_1 - p_0) + x)}{-(q_0 - \overline{\tau} p_0 + x)^2} \\ &= -4 \cdot [\frac{(\overline{\tau} p_1)^2}{(q_0 - \overline{\tau} p_0 + x)^2} + \frac{\overline{\tau} p_1}{(q_0 - \overline{\tau} p_0 + x)}]. \end{aligned}$$

This implies that,

$$R'(x) = -4 \cdot [2 \frac{\overline{\tau} p_1}{q_0 - \overline{\tau} p_0 + x} + 1] \cdot \frac{-\overline{\tau} p_1}{(q_0 - \overline{\tau} p_0 + x)^2},$$

so $R'(x)$ crosses 0 at most once from above as x increases. Therefore, $R(x) \geq R(0)$ on $x \in [0, \delta]$, as $R(0) = R(\delta)$, and $R(x) \geq R(0)$ on $[0, x']$. □

Proof of Theorem 1. If Ω is effectively downward-closed, the proof that some myopic strategy is sequentially optimal is the same as in the proof of Proposition 1. The proof that profits always rise under this strategy is the same as in the proof of Proposition 3.

Suppose Ω is not effectively downward-closed. Let h_t be a reachable history under $\underline{\sigma}$ where for any $p^* \in \arg\max_p p \cdot \underline{d}_{h_t}(p)$ and $d \in \Omega_{h_t}$,

$$p^* \cdot \underline{d}_{h_t}(p^*) < \max_p p \cdot d(p). \tag{5}$$

Let $m \equiv \inf_{d \in \Omega} \max_p p \cdot d(p)$.

We first show that $p^* \cdot \underline{d}_{h_t}(p^*) < m$. Take a sequence of demand curves $d_n \in \Omega$ so that $m - \max_p p \cdot d_n(p) < \frac{1}{n}$. This sequence of demand curves is (1) uniformly bounded, as Ω^{-1} is uniformly equicontinuous, and (2) uniformly equicontinuous, as it is a subset of Ω. Therefore, by the Arzelà-Ascoli theorem, there is a uniformly convergent subsequence of $\{d_n\}_{n\in\mathbb{N}}$. This subsequence converges uniformly to some demand curve $d \in \Omega$, because Ω is closed under the supremum norm, and $\max_p p \cdot d(p) = m$. Therefore, by inequality 5, $p^* \cdot \underline{d}_{h_t}(p^*) < m$.

Finally, because Ω uniformly equicontinuous, there is a fine enough partition of prices such that if the firm experiments with all those prices, it discovers a price where flow revenue is greater than $m + \frac{m - p^* \cdot \underline{d}_{h_t}(p^*)}{2}$. A sufficiently patient would prefer experimenting to discover such a price over using the myopic strategy. □

Lemma **5.** *Prices and profits always rise under σ_{cvx}.*

Proof. Let $d \in \Omega$, and let $h_t = \{(p_1, d(p_1)), \dots (p_t, d(p_t))\}$ be the t-truncation of $\mathbf{h}(\underline{\sigma}, d, h_0)$, for $t \geq 1$. Similarly, let h_{t-1} be the $t-1$ truncation. Let $\underline{\sigma}(h_t) \equiv p_{t+1}$.

At a history $h_1 = \{(p_1, q_1)\}$, $\sigma_{cvx} = \sigma_{lin}$, so by Proposition 1 and Proposition 3, $\sigma_{cvx}(h_1) \geq p_1$ and $\sigma_{cvx}(h_1) \cdot d(\sigma_{cvx}(h_1)) \geq p_1 \cdot d(p_1)$.

Otherwise, suppose h_t is such that $p_t = p_{t-1}$. Then $\sigma_{cvx}(h_t) = p_t$, so flow profit also remains constant.

Finally, suppose that h_t is such that $p_t > p_{t-1} > \dots > p_1$ for $t \geq 2$. It follows from induction that $d_{h_{t-1}}$ lies weakly below the lower envelope of all convex demand curves in Ω on $p \in [p_{t-1}, \infty)$. Then, $\tau \geq \frac{d_{h_{t-1}}(p_t) - d_{h_{t-1}}(p_{t-1})}{p_t - p_{t-1}}$, so marginal revenue (with respect to price) at p_t when demand is given by $d_{h_{t-1}}$ is weakly lower than marginal revenue when demand is d_{h_t}. This implies that $p_{t+1} \geq p_t$. Moreover,

$$p_t \cdot d(p_t) = p_t \cdot d_{h_t}(p_t) \leq p_{t+1} \cdot d_{h_t}(p_{t+1}) \leq p_{t+1} \cdot d(p_{t+1}),$$

where the first inequality is by definition of σ_{cvx} and the second inequality is because d_{h_t} lies weakly below the lower envelope of Ω on $p \in [p_t, \infty)$.

$\square$

Proof of Proposition 7. Parts 2 and 4 follow the same arguments as the baseline case. Part 3 follows from part 2 and lemma 5. We prove part 1 fo histories where the proofs are not redundant.

Case 1: empty history. Suppose $\sigma_{cvx}(h_0) > 0$. Let $f(p) \equiv q_0 + \underline{\tau}(p - p_0)$. Because f is the lower envelope of Ω for $p \geq p_0$, by lemma 5, $\Pi(\sigma_{cvx}, f, h_0) \leq \Pi(\sigma_{cvx}, d, h_0)$ for any $d \in \Omega$. Moreover, $\Pi(\sigma_{cvx}, f, h_0) \geq \Pi(\sigma, f, h_0)$ for any strategy σ by construction.

The case where $\sigma_{cvx}(h_0) < p_0$ is identically proven, with the worst-case state being $g(p) \equiv q_0 + \overline{\tau}(p - p_0)$.

Finally, suppose $\sigma_{cvx}(h_0) = p_0$. This implies that $\arg\max_p p \cdot f(p) \leq p_0 \leq \arg\max_p p \cdot g(p)$. The mapping $\tau \to \arg\max_p q_0 + \tau(p - p_0)$ is singleton-valued and continuous by Berge's theorem. By the intermediate value theorem, there is a $\tau' \in [\underline{\tau}, \overline{\tau}]$ such that $p_0 = \arg\max_p p(q_0 + \tau'(p - p_0))$. Let $l(p) \equiv q_0 + \tau'(p - p_0)$. Because q_0 is a lower bound on demand at p_0 and again lemma 5, $\Pi(\sigma_{cvx}, l, h_0) \leq \Pi(\sigma_{cvx}, d, h_0)$ for any $d \in \Omega$. Therefore, (σ_{cvx}, l) is a saddle point of $\Pi(\cdot, \cdot, h_0)$.

Therefore, σ_{cvx} is ex-ante optimal.

Case 2: $h_1 = \{(p_1, q_1)\}$, where $p_1 < p_0$. Suppose $\sigma_{cvx}(h_1) \equiv p_2 > p_1$. Let $f(p) \equiv \max\{q_0 + \overline{\tau}(p - p_0), q_1 + \underline{\tau}(p - p_1)\}$, and let $[p_1, b]$ be the range of prices p where $f(p) = q_1 + \underline{\tau}(p - p_1)$. Now, for any $p \geq b$, $f(p) = q_0 + \overline{\tau}(p - p_0)$, so

$$p \cdot f(p) < p_1 \cdot (q_0 + \overline{\tau}(p_1 - p_0)) \leq p_1 \cdot q_1 = p_1 f(p_1),$$

which implies that $p_2 \in (p_1, b)$ and $p_2 = \arg\max_p p \cdot f(p)$. Because f is convex, we conclude that $\Pi(\sigma_{cvx}, f, h_1) \geq \Pi(\sigma, f, h_1)$ for any strategy σ. Moreover, f is the lower envelope of Ω on $[p_1, p_0]$, so by lemma 5, $\Pi(\sigma_{cvx}, f, h_1) \leq \Pi(\sigma_{cvx}, d, h_1)$ for any $d \in \Omega$.

If, instead, $p_2 = p_1$, let $\tau' \in [\underline{\tau}, \overline{\tau}]$ be such that $p_2 = \arg\max_p p(q_1 + \tau'(p - p_1))$. Let $f(p) \equiv \{q_0 + \overline{\tau}(p - p_0), q_1 + \tau'(p - p_1)\}$. Now the proof proceeds just as in the previous subcase.

□

B Quantity experimentation

The baseline model studies a firm that sets prices and discovers sales. In settings where firms are capacity constrained or have an inflexible production process, firms may be better thought of as choosing quantity and letting market forces determine prices. Here we consider how quantity experimentation compares to price experimentation.

By symmetry, it is immediate that the form of the sequentially optimal quantity experimentation is identical, after a change of variables. Under this strategy, profits would always rise and the guaranteed ratio of long-run flow profits to informed monopoly flow profits is the same as in the baseline case.

However, there are two key economic differences. First, prices always fall over time, contrary to the direction of prices in Proposition 3. Second, if prices ever strictly fall, then the long-run price under the sequentially optimal quantity experimentation strategy would be *greater* than any informed monopoly price, contrary to the under-pricing in Proposition 4.

The reason for this difference is easy to see. Under sequentially optimal experimentation of either variety, the firm makes the same initial conservative estimate of demand and chooses quantity or price accordingly. If the firm chooses a price p_1, it would typically learn that actual demand, $D(p_1)$, exceeds estimated demand, $d_{h_0}(p_1)$, at that price. This creates an incentive to raise the price. If the firm instead chooses a quantity $x_1 = d_{h_0}(p_1)$, it would learn that its product actually fetches a greater price, $D^{-1}(x_1)$ than the price it had anticipated $d_{h_0}^{-1}(x_1)$. This creates an incentive to raise its quantity, which corresponds to some lower price, as demand curves are downward sloping.

This result on the path of sequentially optimal quantity experimentation squares with the intuition that, say, capacity constrained firms should start small and expand over time if their product sells well.