Title
Learning Through Price Experimentation

Authors
Aghion, Philippe
Bolton, Patrick
Jullien, Bruno

Publication Date
1987-08-03

Peer reviewed

UNIVERSITY OF CALIFORNIA, BERKELEY

Department of Economics

Berkeley, California 94720

Working Paper 8748

LEARNING THROUGH PRICE EXPERIMENTATION BY A MONOPOLIST FACING UNKNOWN DEMAND

Philippe Aghion, Patrick Bolton and Bruno Jullien

August 3, 1987

Key words: Learning by experimentation, adjustment process, Bayesian updating, price-experimentation.

Abstract

This paper investigates the question of how much information firms can acquire about the demand for their product when they learn from experience (i.e. from data about past sales and prices). The main issues are whether firms will eventually learn everything about the demand curve and how learning considerations affect the pricing decisions of firms. It is shown that even when demand is deterministic, strong conditions are required, such as continuity and quasi-concavity of the profit-function, to guarantee that a monopoly will eventually learn all the relevant information about demand.

JEL Classification: 026

ACKNOWLEDGEMENT

We are greatly indebted to Jerry Green, Andreu Mas-Colell, Eric Maskin, Iraj Saniee and Jean Tirole. We also benefited from many helpful comments and discussions with seminar participants at Stanford, Harvard, Chicago, UCLA, Cornell and Berkeley.

Phillipe Aghion
Bruno Jullien

Harvard University

Patrick Bolton

University of California
Berkeley

Section I: Introduction:

In almost all existing theories of market pricing it is assumed that firms know their demand curve. This is true in the theory of monopoly-pricing as well as in most theories of oligopoly pricing. It is such a well-accepted assumption that most economists do not even bother to justify it. Yet one may wonder how firms do acquire the knowledge about their demand curve. In practice firms often conduct market surveys. These give some idea about the profitability of a given market. For instance, they inform the seller about what characteristics of the good consumers like best. They also produce estimates of volume of sales at a given price. However, these surveys do not reveal the entire demand curve to the firm. At best, they allow the firm to find one point on the demand curve.

An alternative source of information for firms is data on current and past sales at the prices set by firms today and in previous periods. Such data is more or less informative depending on the stability of demand over time and mainly on the pricing-rule followed by firms. For example, if the firm never (or rarely) moves its price then the history of past sales will provide a good estimate of expected sales at that price. Or if the firm only allows small price-variations then the history of past prices and outputs will provide a good estimate of the elasticity of demand, at one price. Electric utilities, for example, follow price rules that typically involve small price-variations so that the numerous existing econometric studies of U.S. residential electricity demand (see Bohi (1981) for a survey) at best provide information about price-elasticities at one point on the demand curve. They generally do not provide information about the entire demand curve.

The subject of the present paper is to determine exactly how much a firm can learn from past data, when it sets prices at every period so as to maximize profits. The main issues here are: how do learning considerations affect the firm's pricing decisions and assuming that demand remains stable over time will the firm eventually learn all the relevant information about demand?

The firm's learning process through time can be viewed as an adjustment process towards some equilibrium in the absence of an auctioneer: Suppose that the firm experiences a shock on demand. It will notice this shock, for example, by observing that its inventories are unusually high or low. While it may know that demand has changed in a certain direction, it may not know exactly what the new demand looks like. It will then grope its way towards the new equilibrium by experimenting with prices. Thus the learning problem studied here is relevant to stability theory. There is one major difference, however, with standard stability theory; namely that here it is costly to experiment. The firm foregoes short-run profits by sometimes setting its price too low or too high. If learning is costly, then it may well pay the firm to stop learning before it knows all the relevant information about the new demand.

As a first step towards understanding learning by experimentation, we consider the case of a monopoly. We hope to deal with the oligopoly case in future research. We model the firm as starting with an initial prior distribution over some space of demand functions. These demand functions are assumed to be deterministic. When the firm sets a price, it observes how much it was able to sell at that price and uses this information to update its prior distribution using Bayes' rule. There are two types of firms one can consider in this context: myopic and non-myopic firms. The myopic firms do not understand that by manipulating today's prices, they can gain more or less information about demand tomorrow. They are, however, able to use data about past sales to update their priors. We analyze both types of learning behavior.

The existing literature on this question comprises only a handful of papers: Rothschild (1974); Grossman-Kihlström-Mirman (1977); McLennan (1984); and

Lazear (1986). With the exception of Lazear, all these studies assume that the firm's demand is stochastic. Rothschild (1974) pointed out that when demand is stochastic the firm's learning problem is essentially equivalent to a multi-armed bandit problem (see DeGroot (1970)). Through this analogy, he was able to draw on well known theorems in multi-armed bandit theory, to conclude that a firm will always stop experimenting before it has full knowledge about the demand curve and that as a consequence the firm may end up setting a price different from the full-information monopoly price. In other words, market experimentation does not lead to perfect knowledge of demand.

As Rothschild pointed out, this is an important conclusion since it suggests that existing theories of monopoly pricing cannot be viewed as representations of a firm's pricing behavior, once this firm has decided to stop learning. The long-run equilibrium outcome cannot be separated from the history of price experiments, and from the firm's initial prior information. In other words, when learning is incomplete in equilibrium, we have a situation where the long-run equilibrium is <u>history-dependent</u>. The significance of this conclusion is that standard comparative statics analyses have to be abandoned when learning is incomplete and history becomes a fundamental explanatory variable.

In a subsequent study, McLennan also obtained that when the demand function is stochastic firms will not have all the relevant information about demand in long-run equilibrium. Thus both of these studies seem to indicate that the incomplete-learning results stem from the stochastic nature of demand. Indeed, if one looks at the deterministic version of the demand-function in McLennan then the firm will end up with complete knowledge of demand in long-run equilibrium.

Our study assumes that the firm faces the same deterministic demand every period, and we show that Rothschild's incomplete-learning result does not depend on the stochastic nature of demand. Assuming a deterministic demand simplifies the analysis considerably. The firm's inference problem is now trivial (the firm learns a point on the demand curve whenever it sets a price). Simplifying the inference problem allows us to characterize the firm's learning strategy much more precisely. We are able to substantially generalize Rothschild's and McLennan's results and to identify the reasons for which a monopoly may want to stop learning before knowing the true demand curve. In addition, a simple methodology is developed here to handle adaptive-learning problems in deterministic settings. This methodology may be applicable to a wide range of problems.

The paper is organized as follows: Sections II and III set out two polar cases of imperfect knowledge about a discontinuous deterministic demand, which can only be learned through price experimentation. In the first case the monopoly price can take a finite number of values but the quantities sold at each price are uncertain. In the second case, all the consumers have the same reservation value, which is unknown. Section III develops a useful analytical tool for learning through experimentation problems, which is used in Section IV. This last section discusses the extension of the results obtained in Sections II and III to arbitrary continuous demand curves. It is argued that in the case where the demand curve is known to be continuous, the monopoly will eventually set a price arbitrarily close to a local optimum of the profit function but may never learn the global optimum. Finally, Section V offers some concluding comments.

Section II: The case where the consumers' reservation values are distributed on a finite grid of prices.

In this first model, we assume that there are two types of consumers: those who attach a high value to the monopolist's product and those who attach a low value. Each consumer purchases at most one unit of the good. Thus some consumers have a reservation value v_1 and others a reservation value v_2, where $v_1 > v_2$.

The proportion of consumers with a high reservation value is given by $\mu \in [0,1]$. The consumers are infinitely lived and have the same demand for the good each period. The monopolist produces the good at unit cost, $c=0$, and has a discount factor $\delta \in (0,1)$. Every period he can set a new price. If one period is a day, for example, then the monopolist must charge the same price to every customer during that day. (One can imagine a situation where the firm advertises its prices in the newspapers and thus can only change its price by changing the advertisement.)

If the firm has complete information about demand it will set its prices as follows:

$$P_t = v_1 \iff \mu \cdot v_1 \geq v_2$$
$$(1)$$
$$P_t = v_2 \iff \mu \cdot v_1 < v_2$$

Thus, when the firm has complete information about demand, the price P_t is constant through time. Now suppose that v_1 and v_2 are known to the firm, but that μ is unknown. Then, if the firm sets the initial price $P_1 = v_2$, it will learn nothing about μ. Hence, if it is optimal to set $P_1 = v_2$, it must also be optimal to set $P_t = v_2$, for all $t \geq 1$.

Next, if the monopolist sets $P_1 = v_1$, he will learn immediately the true value of μ and he will set P_t as in (1) for all $t \geq 2$.

Let $f(\mu)$ be the prior distribution of μ on $[0,1]$, then the expected net present value of profits for the monopolist, when he sets $P_1 = v_1$, is given by:

$$(2) \quad \Pi_1 = \int_0^{\bar{\mu}} (\mu v_1 + \frac{\delta v_2}{1-\delta}) f(\mu) d\mu + \int_{\bar{\mu}}^1 (\mu v_1 + \frac{\delta \mu v_1}{1-\delta}) f(\mu) d\mu$$

where $\bar{\mu} = \frac{v_2}{v_1}$.

And it will be optimal to set $P_t = v_2$, for all t, when

$$(3) \qquad \frac{v_2}{1-\delta} \geq \Pi_1$$

where the LHS of (3) is the net present value of profits, when $P_1 = v_2$ is the optimal initial price. Now suppose, for example, that $f(\mu)$ is uniform on $[0,1]$, then we obtain:

$$(4) \qquad \Pi_1 = \frac{v_1}{2(1-\delta)} + \frac{\delta v_2^2}{2v_1(1-\delta)} .$$

so that $\frac{v_2}{1-\delta} \geq \Pi_1$ is equivalent to

$$(5) \qquad \frac{v_1}{v_2} \leq 1 + \sqrt{1-\delta} .$$

This model is very simple, but rather instructive. Firstly, note that the monopolist may choose to set $P_t = v_2$ for all t, while the true proportion μ may be such that $\mu v_1 > v_2$. That is, the monopolist may set the wrong price forever. Thus, in this very simple model we obtain similar conclusions to those of Rothschild (1974) and McLennan (1984). The main difference between our model and, for example, Rothschild's model is that here when the firm learns something about the demand, it learns everything. The inference problem in the context of a deterministic demand is trivial. In Rothschild's model, on the other hand, the firm's inference problem is complicated because demand is stochastic. He shows that even if the firm ends up setting the right price, it will never be certain of this fact. In our model, this is only true if the firm sets $P_t = v_2$. If the firm sets $P_t = v_1$ for some t , then it learns everything about demand and it will know that it sets the correct price.

Secondly, we obtain the Grossman-Kihlstrom-Mirman (1977) result, that learning considerations tend to push up prices. If the firm were only to set one price forever, then it would choose

$$
(6) \qquad \begin{aligned} P &= v_1 \iff \frac{v_1}{2} \geq v_2 \\ P &= v_2 \quad \text{otherwise} \end{aligned}
$$

(μ being uniformly distributed on $[0,1]$, the mean proportion is given by $\bar{\mu} = 1/2$). However, if the firm takes into account learning considerations it will set:

$$
(7) \qquad P_1 = v_1 \iff \frac{v_1}{1 + \sqrt{1-\delta}} \geq v_2 \quad .
$$

Thus, when the firm wishes to learn about its demand curve (i.e., $\delta > 0$) it tends to set higher prices.

Finally, we also obtain the result that prices are either constant through time or decreasing. This is akin to a result obtained by Lazear (1986) in a two-period model.

Do these results extend to more general models where there are n possible reservation values, $(v_1, v_2, \ldots, v_n)$ and where demand is not necessarily decreasing everywhere? Unfortunately, only the result of incomplete learning is robust. As soon as we have more than two reservation values, the other two results are in general no longer valid. The following two examples illustrate the possibility of having on the one hand, prices increasing over time and on the other hand, the initial price being set below the price that equates marginal cost with expected marginal revenue (in other words, learning considerations pull prices down).

Example 1: Suppose that there are three possible reservation values, $v_1 > v_2 > v_3$, and that μ_1, μ_2, μ_3 are respectively the proportion of consumers with reservation values, v_1 , v_2 and v_3 ($\mu_i \in [0,1]$; $i = 1,2,3$ and $\sum_{i=1}^{3} \mu_i = 1$). Each proportion μ_i can take two possible values. These are given in the table below.

	μ_1	μ_2	μ_3
α	2/3	0	1/3
$(1-\alpha)$	0	1/3	2/3

(where $\alpha = \Pr(\mu_1 = 2/3 ; \mu_2 = 0 ; \mu_3 = 1/3)$. Thus μ_1 can take the value 2/3 with probability α and zero with probability $(1-\alpha)$. The monopolist does not know in advance what the proportion of consumers with a particular reservation value is, but he knows α .

We shall assume that $3v_3 > v_2$ and $2v_1 > 3v_3$, so that if $\alpha = 1$, the optimal price is v_1, and if $\alpha = 0$, the optimal price is v_3. When $\alpha \varepsilon (0,1)$, the monopolist can learn the true proportions by setting the initial price P_1 equal to v_2. His intertemporal profits then are:

$$\Pi_2 = \alpha[\frac{2v_2}{3} + \frac{2v_1}{3(1-\delta)}] + (1-\alpha)[\frac{v_2}{3} + \frac{v_3}{1-\delta}] .$$

If he sets $P_1 = v_3$, he learns nothing and his intertemporal profits are given by:

$$\Pi_3 = \frac{v_3}{1-\delta}$$

If he sets $P_1 = v_1$ he also learns the true proportions and his intertemporal profits are:

$$\Pi_1 = \alpha(\frac{2v_1}{3(1-\delta)}) + (1-\alpha)(\frac{\delta v_3}{1-\delta})$$

We then have prices increasing from v_2 to v_1 with probability α whenever $\Pi_2 > \Pi_1$ and $\Pi_2 > \Pi_3$. That is, when:

(8) $$\alpha < \frac{v_2}{2v_1 - v_2}$$

(9) $$\alpha > \frac{(3v_3 - v_2)(1-\delta)}{v_2 + \delta(2v_1 - v_2 - 3v_3)} .$$

The inequalities (8) and (9) are compatible when

(10) $$\frac{v_2}{2v_1 - v_2} > \frac{(3v_3 - v_2)(1-\delta)}{v_2 + (2v_1 - v_2 - 3v_3)} .$$

And (10) is always verified when δ is large or when v_1 is small. Specifically when

(11) $$2v_1 < \frac{3v_3v_2}{3v_3-v_2} .$$

This example illustrates the possibility that the monopolist will be cautious in experimenting with prices and will only raise prices as he becomes more optimistic about demand.

Example 2: Suppose now that there are four possible reservation values $v_1 > v_2 > v_3 > v_4$ and that the proportions μ_1,μ_2,μ_3,μ_4 are given as follows:

	μ_1	μ_2	μ_3	μ_4
α	2/3	0	1/3	0
$1-\alpha$	0	2/3	0	1/3

We assume that: $\frac{2v_2}{3} > \alpha v_3 + (1-\alpha)\frac{2v_3}{3}$ and that $\frac{2v_2}{3} > \alpha\frac{2v_1}{3}$.

Then, if the monopolist can only set one price forever, he would choose $P = v_2$. We will show, however, that if he can set a different price each period, it is possible that learning considerations will induce him to first set $P_1 = v_3$ and then $P_t = v_1$ if it turns out that $\mu_1 + \mu_2 + \mu_3 = 1$. Otherwise he will set $P_t = v_2$ for all $t \geq 2$. If $P_1 = v_3$ the monopolist's expected intertemporal profits are:

$$\Pi_3 = \alpha[v_3 + \frac{\delta 2v_1}{3(1-\delta)}] + (1-\alpha)\ [\frac{2v_3}{3} + \frac{\delta 2v_2}{3(1-\delta)}] .$$

If $P_1 = v_1$,

$$\Pi_1 = \frac{\alpha 2v_1}{3(1-\delta)} + (1-\alpha)\ \frac{\delta 2v_2}{3(1-\delta)} \quad .$$

Finally, v_2 and v_4 are completely uniformative prices, so that

$$\Pi_2 = \frac{2v_2}{3(1-\delta)} \quad \text{and} \quad \Pi_4 = \frac{v_4}{(1-\delta)} \quad .$$

Now, $P_1 = v_3$ is optimal if and only if $\Pi_3 > \Pi_i$ for $i = 1,2,4$. And, $\Pi_3 > \Pi_1 <=> \alpha < \frac{2v_3}{2v_1 - v_3}$. Assume that $\frac{2}{3} v_2 > v_4$ so that $\Pi_2 > \Pi_4$, then

$$\Pi_3 > \Pi_2 \quad <=> \quad \alpha > \frac{\frac{2}{3}(v_2 - v_3)}{\frac{v_3}{3} + \frac{2v_1}{3(1-\delta)} - \frac{\delta 2v_2}{3(1-\delta)}} \quad .$$

Therefore, for α small enough and δ large enough, all the inequalities are verified and $P_1 = v_3$ is optimal. Thus, this example shows that the principle that learning considerations tend to push up prices is not valid in general.[1]

Section III: The case where all consumers have the same reservation value. Model II.

In the preceding section, we developed a model where the reservation prices were known to the monopolist, but not the proportion of consumers corresponding to each reservation value. Here we will analyze a polar case, where all consumers have the same reservation value, but where the monopolist does not know this value. Thus, when he sets a price, the monopolist either serves the entire market and learns that the reservation value is above this price or he does not serve any customer and learns that the reservation value is below the price he set.

We shall consider two types of firms: myopic and non-myopic price setters. The former learn from experience but do not understand that today's experiment can be manipulated so as to acquire more or less information tomorrow. The latter are aware of the effect of today's pricing decision on tomorrow's value of information.

Assume that the consumers' reservation value v can be any number in the interval $[0,1]$. As in the previous section, consumers purchase at most one unit every period and do not behave strategically (i.e., they buy whenever $P_t \leq v$). The firm produces the good at unit cost $c=0$ and has a discount factor $\delta \in (0,1)$. The monopolist's prior distribution on v is assumed to be uniform on $[0,1]$.[2]

This is essentially the same as Lazear's model except that here we have repeat purchase while in Lazear (1986), consumers buy the good only once. It turns out that allowing for repeat purchase changes the conclusions substantially. While the above model is overly simplistic, it can be viewed as a representation of markets where consumers' tastes are very similar and where each consumer has a unit demand. Surprisingly, even within such a simple framework the firm's dynamic pricing policy is very difficult to characterize.

We begin with the decision problem faced by a <u>non-myopic</u> monopolist: For any pair (x,y) where $0 \leq x < y \leq 1$, we denote by $V_\delta(x,y)$ the maximum expected intertemporal profits when the firm's initial information is that v is uniform on $[x,y]$. The optimization behavior of the non-myopic firm can then be described as follows: At time $t = 0$, it chooses some price $P_0 \varepsilon [0,1] = I_0$. Since all consumers are identical, there are only two possible outcomes: Either the firm serves the entire market at price P_0 (we denote this by $\mu^0 = 1$, where μ^0 is the volume of sales at time $t = 0$), or the firm does not sell to anyone (this is denoted by $\mu^0 = 0$). If $\mu^0 = 1$, the monopolist can update his information about v to $v \varepsilon [P_0,1] = I_1$. Similarly, if $\mu^0 = 0$, the monopolist's information about v becomes: $v \varepsilon [0,P_0] = I_1$. Now, the optimal choice of the initial price P_0 made by a non-myopic monopolist is the solution to the following maximization program:

$$P_0^* = \arg \max_{P_0 \varepsilon [0,1]} [(1-P_0)P_0 + \delta(1-P_0)\, V_\delta(P_0,1) + \delta P_0 \cdot V_\delta(0,P_0)]$$

Thus, the characterization of the optimal price schedule followed by a non-myopic firm involves essentially the study of the valuation function V_δ. This is the main difference with the case of a myopic firm. The latter simply chooses P_0 to maximize:

$$\max_{P_0 \varepsilon [0,1]} \quad Pr(\mu^0 = 1/P_0) \cdot P_0 = (1-P_0)P_0 \quad .$$

As in most dynamic programming problems, it is useful to redefine V_δ as the solution to a Bellman equation. This equation will enable us to derive important properties of this valuation function and to characterize the learning process followed by the firm.

<u>Proposition 1</u>: $V_\delta(x,y)$ is the unique bounded solution of the following Bellman equation:

$$(B): \quad V(x,y) = \max_{z\varepsilon[x,y]} 1/y-x \,\{(y-z)z + \delta\cdot(y-z)V(z,y) + \delta(z-x)V(x,z)\} .$$

<u>Proof</u>: Suppose that the monopolist starts with the information that v is uniformly distributed on $[x,y]$. When he sets the first-period price, $z\varepsilon[x,y]$, he will know at the end of period 1 whether $v\varepsilon[x,z]$ or $v\varepsilon[z,y]$, and his expected profits will be:

$$\Pi(z,x,y) = z\cdot\psi_1(z,x,y) + \delta\psi_1(z,x,y)V_\delta(z,y) + \delta(1-\psi_1(z,x,y))\cdot V_\delta(x,z)$$

$$= z \cdot\left(\frac{y-z}{z-x}\right) + \delta\left(\frac{y-z}{y-x}\right)V_\delta(z,y) + \delta\left(\frac{z-x}{y-x}\right)V_\delta(x,z)$$

Now, $V_\delta(x,y)$ is the maximum of $\Pi(z,x,y)$ over $z \in [x,y]$. This shows that V_δ is a solution of the Bellman equation (B). Next, we show that V_δ is actually the unique bounded solution of this equation. $V_\delta(x,y)$ is bounded since it can be written as an infinite sum: $\Sigma\delta^t\cdot\Pi_t(x,y)$, where $0 \leq \Pi_t(x,y) \leq 1$ for all t and all pairs (x,y) . ($\Pi_t(x,y)$ is the maximum expected profit at time t given the prior information at time 0: $v \in [x,y]$.)

Now, let B_0 denote the space of bounded functions on $[0,1]^2$. B_0 is a Banach space for the uniform norm $|\cdot|_\infty$ defined by: $|f|_\infty = \sup_{x\varepsilon[0,1]^2}|f(x)| < +\infty$ for all $f \in B_0$. Next, consider the following mapping Π :

$$\Pi: \quad B_0 \to B_0$$
$$V \mapsto \Pi(V)$$

where:

(1) $$\Pi(V)_{(x,y)} = \max_{z \varepsilon [x,y]} \left\{ \frac{y-z}{y-x} z + \delta \frac{y-z}{y-x} V(z,y) + \delta \frac{z-x}{y-x} V(x,z) \right\}$$
$$= g_V(x,y,z) \quad .$$

We know that $V_\delta \; \varepsilon \; B_0$ is a solution of the Bellman equation (B) (i.e., a fixed point of the mapping Π). The following lemma establishes that such a fixed point is unique on B_0 .

Lemma: Π is a contraction mapping on B_0 .

Proof of the lemma: Let V and $V' \; \varepsilon \; B_0$, and

$$z_v(x,y) = \arg\max_{z \varepsilon [x,y]} g_v(x,y,z) \quad .$$

We have:

$$|\Pi(V)_{(x,y)} - \Pi(V')(x,y)| = |g_v(x,y,z_v(x,y)) - g_{v'}(x,y,z_{v'}(x,y))|$$
$$\leq g_v(x,y,z_v(x,y)) - g_{v'}(x,y,z_v(x,y))$$

(whenever $g_v(x,y,z_v(x,y)) \geq g_{v'}(x,y,z_{v'}(x,y))$, which we can assume w.l.o.g.).

Thus:

$$|\Pi(V)_{(x,y)} - \Pi(V')_{(x,y)}| \leq \delta \cdot \left\{ \frac{z_V(x,y)-x}{y-x} [V(x,z_v(x,y)) - V'(x,z_v(x,y))] + \frac{y-z_v(x,y)}{y-x} [V(z_v(x,y),y) - V'(z_v(x,y),y)] \right\}$$

$$\leq \delta \cdot |V - V'|_\infty \cdot \left[\frac{z_v(x,y)-x+y-z_v(x,y)}{y-x}\right]$$

$$\leq \delta \cdot |V - V'|_\infty \quad .$$

Since $\delta < 1$, Π is a contraction mapping on B_0 , which proves the lemma. Now, given that B_0 is a Banach space, Π has a unique fixed point on B_0 , namely V_δ . This completes the proof of Proposition 1. □

V_δ as the unique fixed point of Π can be approximated arbitrarily closely by using the fact that:

$$V_\delta = \lim_{n \to +\infty} \Pi^n(f)$$

(where: 1) f is any element of B_0 ; 2) Π^n is the n-iterate of Π ; 3) the limit is taken w.r.t. the $|\cdot|_\infty$ norm on B_0 .)

Using such an iteration procedure we have been able through simulations to map $V_\delta(x,1)$ as a function of $x \in [0,1]$. The figures below represent $V_\delta(x,1)$ for three different values of δ .

Insert Figure 1 about here.

Corollary: $V_\delta(x,y)$ is homogeneous of degree 1, convex and increasing in x and y .

Proof: Given that V_δ is the unique fixed point on B_0 of the contraction mapping Π defined by (1) above (Proposition 1), it suffices, in order to prove this corollary, to show that for all $V \in B_0$ where V is homogeneous of degree 1, convex, and increasing in x and y , that $\Pi(V)$ inherits the same properties. So, let $V \in B_0$ verify these three properties:

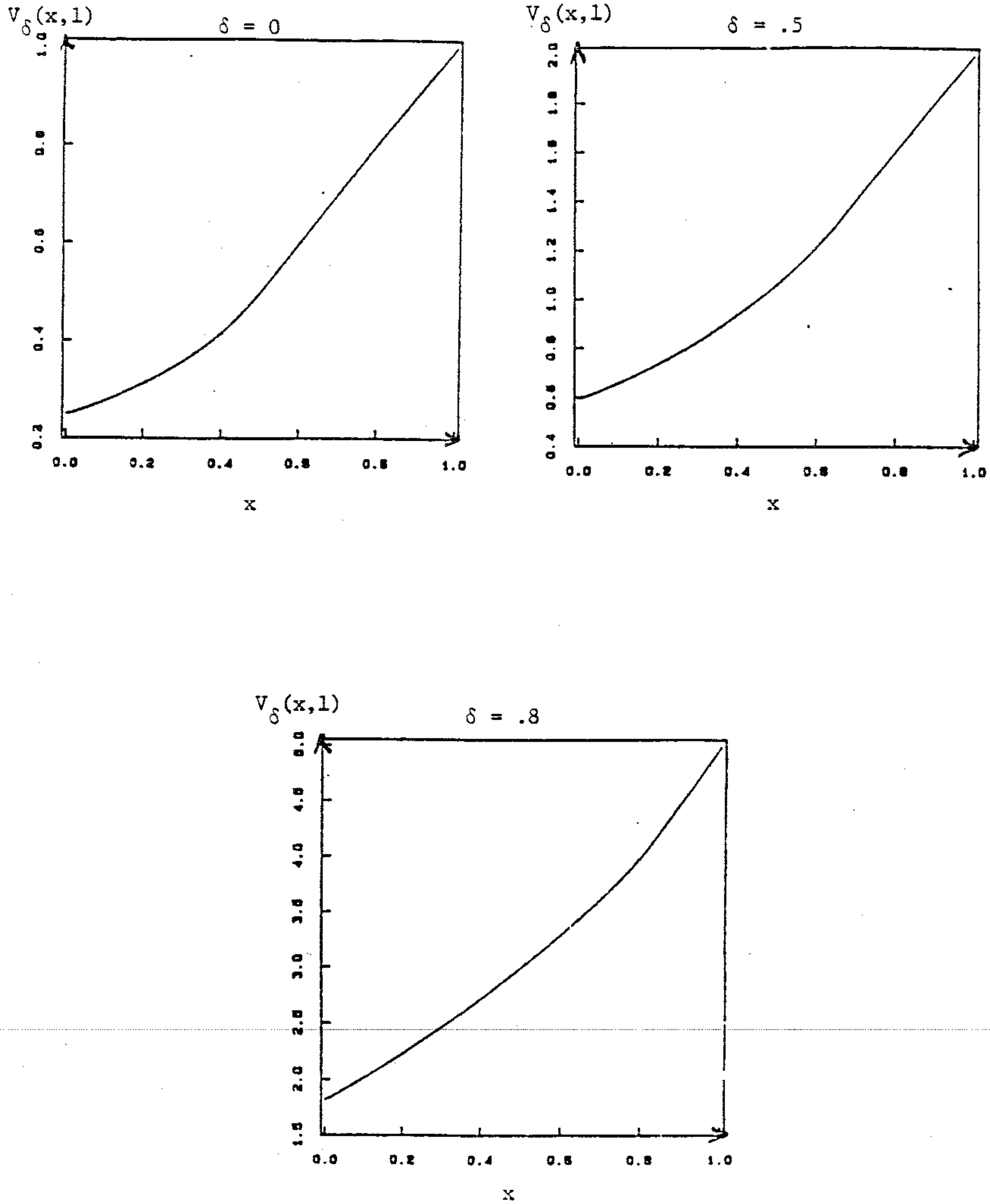

Figure 1

(a) $\Pi(V)$ is homogeneous of degree one.

$$\Pi(V)(\lambda x,\lambda y) = \max_{z\varepsilon[\lambda x,\lambda y]} \{\frac{\lambda y-z}{\lambda(y-x)}\cdot z + \delta\cdot \frac{\lambda y-z}{\lambda(y-x)}\cdot V(z,\lambda y) + \delta\cdot\frac{z-\lambda x}{\lambda(y-x)} V(\lambda x,z)\}$$

$$= \max_{z'\varepsilon[x,y]} \{\frac{y-z'}{y-x}\cdot\lambda\cdot z' + \delta\cdot\frac{y-z'}{y-x} V(\lambda z',\lambda y) + \delta\cdot\frac{z'-x}{y-x} V(\lambda x,\lambda z')\} \text{ , where } z' = \frac{z}{\lambda} .$$

$$= \lambda\cdot\Pi(V)_{(x,y)} \text{ , by homogeneity of degree 1 of } V .$$

(b) $\Pi(V)$ is convex.

Using the fact that any $z \varepsilon [x,y]$ can be written as $z = (y-x)q + x$, where $q \varepsilon[0,1]$, we can express $\Pi(V)$ as:

$$\Pi(V)_{(x,y)} = \max_{q\varepsilon[0,1]} \{h(q,x,y) = (1-q)\cdot((y-x)q+x) + \delta(1-q)\cdot V((y-x)q+x,y) + \delta q\cdot V(x,(y-x)q+x)\} .$$

Given that $V \varepsilon B_0$ is convex, and that $(x,y)\longmapsto (y-x)q+x$ defines a linear mapping on $[0,1]^2$ for each $q \varepsilon[0,1]$, we have:

(a) for each $q \varepsilon[0,1]$, $h(q,x,y)$ is convex in x and y , and

(b) for each $(x,y)\varepsilon[0,1]^2$, $h(q,x,y)$ is continuous in q .

Thus, $\Pi(V)$ is the upper envelope of a continuous family of convex mappings so that $\Pi(V)$ must also be convex in x and y .

(c) $\Pi(V)$ is increasing in x and y .

For each $q \varepsilon [0,1]$, $(x,y)\longmapsto (y-x)q+x = qy + (1-q)x$ is increasing in x and y since $0 \leq q \leq 1$. Given that V is assumed to be increasing in x and y , for each $q \varepsilon [0,1]$, $h(q,x,y)$ increases with x and y and also $\Pi(V)_{(x,y)} = \max_{q\varepsilon[0,1]} h(q,x,y)$.

This establishes the corollary. □

Proposition 1 and its corollary provide us with the main analytical tools for solving our problem. To see how powerful these tools are, it is instructive to contrast the case of a non-myopic firm with the myopic case.

We know that a myopic price setter chooses his initial price P_0 so as to maximize his short-run expected profit: $(1-P_0)P_0$, i.e., he chooses $P_0^* = 1/2$. Then, if $\mu^0 = 1$, i.e., if the firm sells at price P_0^*, it will learn that $v \in [1/2,1] = I_1$. In this case the myopic firm will choose P_1 at time $t = 1$ so as to maximize:

$$\begin{aligned} \Pi_1(P_1,I_1) &= Pr(\mu^1 = 1/P_1 \text{ and } v \in I_1) \cdot P_1 \\ &= Pr(v \geq P_1 / v \in I_1 = [1/2,1]) \cdot P_1 \\ &= \frac{(1-P_1)}{1-1/2} \cdot P_1 = 2(1-P_1) \cdot P_1 \quad . \end{aligned}$$

Therefore, when $\mu^0 = 1$ (i.e., $v \geq 1/2$), the myopic firm will again choose its optimal price P_1^* equal to 1/2. But this price is uninformative since we already know that v is greater than 1/2. This means that a myopic monopolist who sells at time $t = 0$ will fix $P_t^* = 1/2$ for all $t \geq 1$. In other words, it stops experimenting from period 1 on.

What happens now if the monopolist is non-myopic? Would it stop experimenting once it is known that $v \geq 1/2$ (or more generally, by homogeneity of degree one, when $v \in [x,y]$ with $x/y \geq 1/2$)? This question, among others, can be answered by making use of Proposition 1 and its corollary.

Proposition 2: For $\delta \in (0,1)$, let $x_\delta = 1/2-\delta$, then:

(a) $V_\delta(x,1) = x/1-\delta$ if and only if $x \in [x_\delta,1]$.

In other words, whenever $x \geq x_\delta$, the optimal policy for a non-myopic monopolist whose prior information is $v \in [x,1]$ is to play $P_t = x$

forever and not learn anything more on v . However, for $0 \leq x < x_\delta$, the non-myopic monopolist will optimally set its first-period price P_0 strictly <u>above</u> x .

(b) $V_\delta(0,1) > M_\delta(0,1)$ for $\delta > 0$, where

$M_\delta(0,1) = \frac{1}{(4-\delta)(1-\delta)}$ is the maximum expected intertemporal profit of a myopic monopolist with prior information: $v \varepsilon [0,1]$.

<u>Proof</u>: Let $W_\delta(x) = V_\delta(x,1)$ for all $x \varepsilon [0,1]$. (By homogeneity of degree one of V_δ , we have:

$V_\delta(x,y) = y \cdot W_\delta(x/y)$, for all $0 \leq x < y \leq 1$.)

(a) Let z be the smallest x such that $W_\delta(y) = y/1-\delta$ for all $y \varepsilon [x,1]$. Then, from the Bellman equation (B), we must have:

$$W_\delta(z) = \max_{y \varepsilon [z,1]} \frac{1}{1-z} \{y(1-y) + \delta(1-y)W_\delta(y) + \delta(y-z)yW_\delta(\tfrac{z}{y})\}$$

$$= \max_{y \varepsilon [z,1]} \frac{1}{1-z} \{ y(1-y) + \frac{\delta(1-y)y}{1-\delta} + \frac{\delta(y-z)z}{1-\delta} \} = \max_{y \varepsilon [z,1]} G(y)$$

since $z/y \geq z$ and $y \geq z$ and by definition of z . The first order condition for the above maximization is:

$1 - 2y + \delta z = 0$,

i.e., $y = \frac{1+\delta z}{2}$.

Now, $y>z \iff z < \frac{1}{2-\delta} = x_\delta$,

So, for $z \geq x_\delta$, we must have:

$y^* = \arg\max G(y) = z$, i.e.:

$W_\delta(z) = G(z) = \frac{z}{1-\delta}$.

This proves that $V(x,1) = x/1-\delta$ is the solution of the Bellman equation (B) on the set $\{x \varepsilon [0,1] / x \geq x_\delta\}$, and that x_δ is less than or equal to the smallest x such that $V_\delta(y,1) = y/1-\delta$ for all $y \in [x,1]$. Since we know from Proposition 2 that V_δ is the unique bounded solution of this Bellman equation on [0,1] , we necessarily have:

$$V_\delta(x,1) = \frac{x}{1-\delta} \text{ for all } x \in [x_\delta,1] \quad ,$$

Insert Figure 2 about here

and $V_\delta(x,1) \neq x/1-\delta$ for $x < x_\delta$, x close to x_δ .

But since V_δ is <u>convex</u> (Proposition 1), there cannot be any other $x < x_\delta$ such that: $V_\delta(x,1) = x/1-\delta$. (See Figure 2 above.)

This establishes (a).

(b) Take the case of a myopic monopolist starting with the prior information: $v \in [0,1]$. We know that such a monopolist will choose to set the price $P^*_0 = 1/2$ at time $t = 0$. And, if the firm sells at that price, it learns that $v \in [1/2,1]$. But then, our myopic agent sets $P^*_t = 1/2$ forever, i.e.:

$$M_\delta\left(\frac{1}{2}, 1\right) = \frac{\frac{1}{2}}{1-\delta} \quad .$$

In the other case, $v \in [0,1/2]$, the monopolist faces the same problem as if $v \in [0,1]$, by homogeneity of degree one. Thus, the maximum intertemporal profit he can expect is:

$$M_\delta\left(0, \frac{1}{2}\right) = \frac{1}{2} \cdot M_\delta(0,1) \quad .$$

We have:

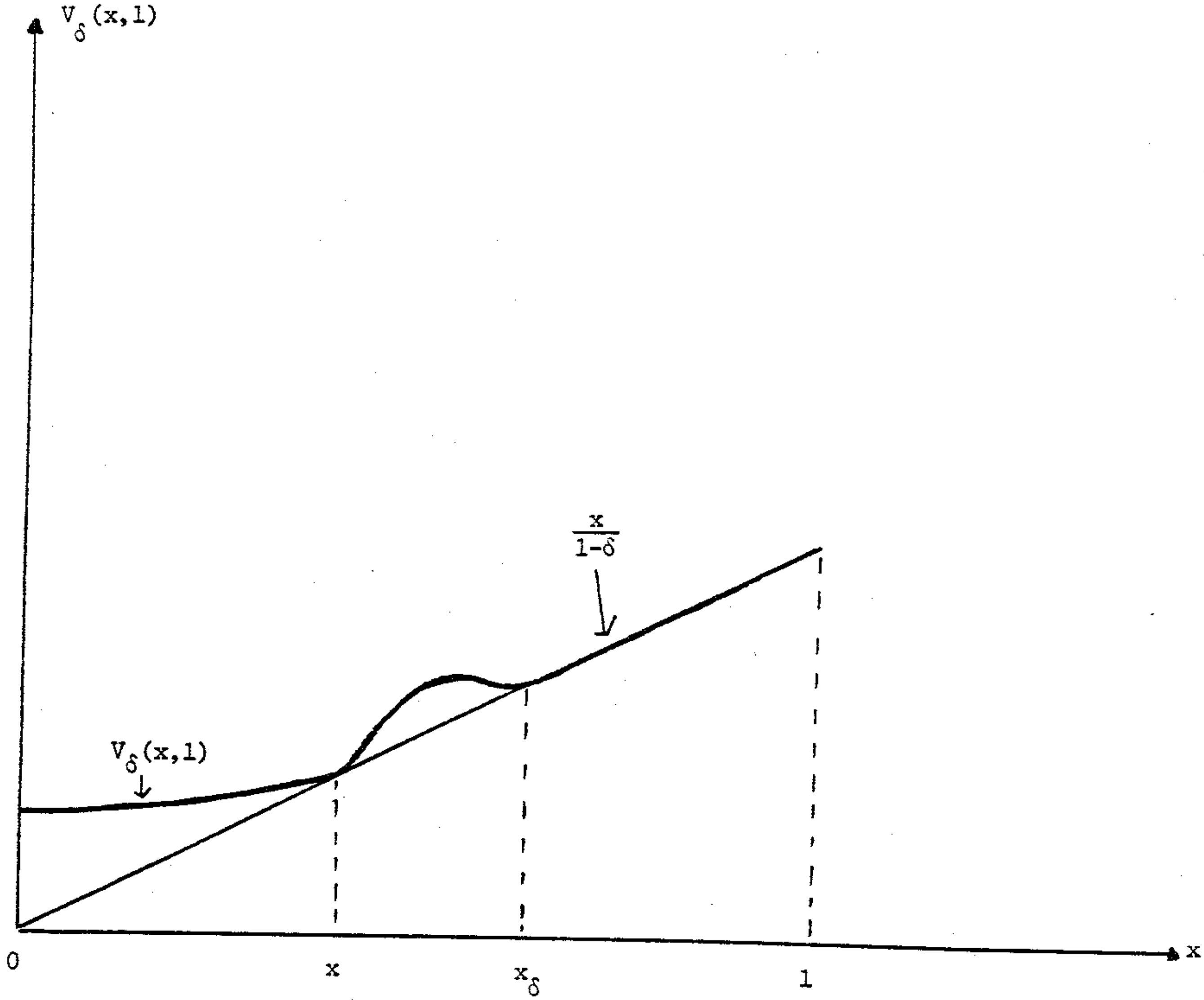

Figure 2

A hump-shaped $V_\delta(x,1)$, as in Figure 2, is impossible.

$$M_\delta(0,1) = \frac{1}{2} \cdot \left(\frac{1}{2} + \delta M_\delta\left(\frac{1}{2} , 1 \right)\right) + \frac{1}{2}\, \delta M_\delta\left(0, \frac{1}{2}\right)$$

$$= \frac{1}{4} + \frac{\delta}{4(1-\delta)} + \frac{1}{4}\, \delta M_\delta(0,1) \quad .$$

$$\Rightarrow \qquad M_\delta(0,1) \cdot (4-\delta) = \frac{1}{1-\delta}$$

$$\Rightarrow \qquad M_\delta(0,1) = \frac{1}{(4-\delta)(1-\delta)} \quad .$$

In the non-myopic case we have the following inequality:

$$V_\delta(0,1) \geq \frac{1}{4} + \frac{\delta}{2}\, V_\delta\left(\frac{1}{2} , 1 \right) + \frac{\delta}{2}\, V_\delta\left(0, \frac{1}{2}\right)$$

where the RHS represents the non-myopic firm's profit when $P_0 = 1/2$.

By homogeneity we have:

$$V_\delta\left(0, \frac{1}{2}\right) = \frac{1}{2} \cdot V_\delta(0,1) \quad ,$$

for $\delta > 0$, $1/2$ is strictly less than $x_\delta = 1/2-\delta$. Therefore, from Proposition 3(a) and from the convexity of V_δ :

$$V_\delta\left(\frac{1}{2} , 1\right) = W_\delta\left(\frac{1}{2} \right) > \frac{\frac{1}{2}}{1-\delta} \quad .$$

Therefore:

$$V_\delta(0,1) > \frac{1}{4} + \frac{\delta}{4}\, \frac{1}{1-\delta} + \frac{\delta}{4}\, V_\delta(0,1)$$

$$\text{i.e.,} \quad V_\delta(0,1) > \frac{1}{(4-\delta)(1-\delta)} = M_\delta(0,1)$$

Proposition 2 is proved. □

Thus a non-myopic firm continues experimenting, unless its information about v is such that $v \in [x_\delta, 1]$. Notice that when $\delta = 0$ this corresponds to the myopic solution and when $\delta = 1$, the firm experiments until it has all the information about demand.

An important consequence of Proposition 2 (the main result of this section), is that unless $\delta = 1$, the non-myopic firm (and a fortiori the myopic firm) never ends up with perfect information about demand. As a result it cannot be certain that it is setting the best possible price.

Theorem 1:

(a) Except when the true value v is zero, the sequence $(P_t(v))$ is eventually non-decreasing.

(b) Except when the true value v is zero, the nested sequence of information sets $I_t = [\underline{v}_t , \bar{v}_t]$ will never converge to the singleton $\{v\}$. The length $\ell(I_t)$ will remain bounded away from 0 when $t \to \infty$.

(Part b) of the theorem is represented in the diagram below:)

Insert Figure 3 about here.

Although the sequence (P_t) may converge to the true value $v = \underline{v}_\infty$, the monopolist's information converges to $I_\infty = [v,\bar{v}]$, where $\bar{v} > v$; and the monopolist does not know that he is setting the correct price.

Proof:

(b) Suppose that the true reservation value v is strictly positive, and suppose that the sequence of information sets (I_t) (where $I_0 = [0,1]$) converges to the singleton $\{v\}$. Then, for $t \geq T$ (T large enough), $I_t = [\underline{v}_t , \bar{v}_t]$ must be such that: $\underline{v}_t/\bar{v}_t \geq x_\delta = 1/2-\delta$. But then, by homogeneity of degree one of the valuation function V_δ , and from Proposition 3(a), we know that the monopolist will set: $P_t = \underline{v}_t$ for all

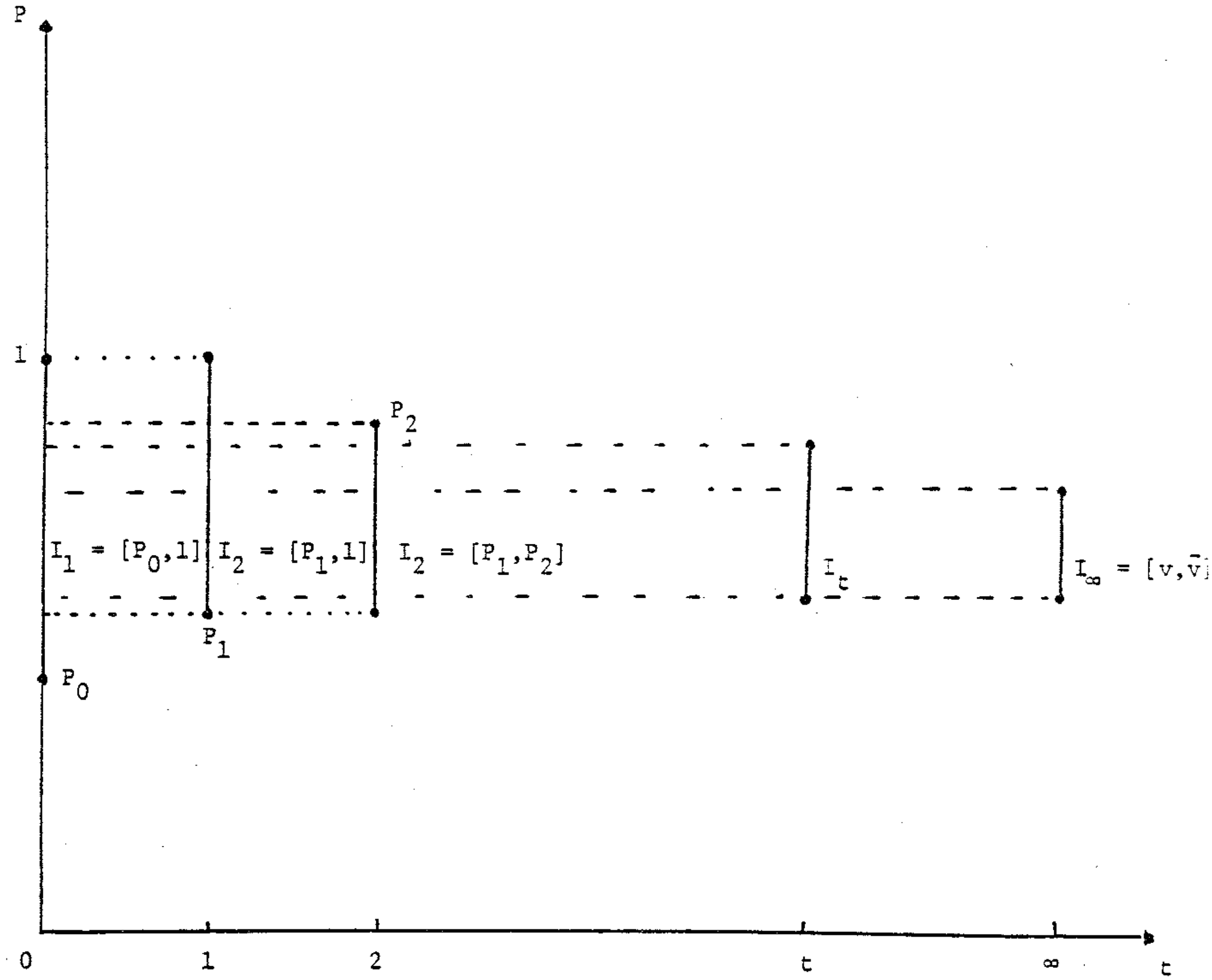

Figure 3

$t \geq T$. This implies that $I_t = I_T$ for all $t \geq T$, a contradiction of the assumption that $I_t \to \{v\}$.

(a) Assume that $v \neq 0$ and suppose that (a) is not satisfied for some price sequence $P_t(v)$. Then we could always extract a subsequence of P_t which is decreasing and bounded below by $v > 0$ (if $P_t < v$ for some t , then necessarily $P_{t+1} \geq P_t$ and P_t would not be decreasing!). Therefore the sequence P_t should converge from above to some accumulation point $w \geq v > 0$. This means in particular that for t large enough, the price P_{t+1} set at time $t+1$ should be chosen arbitrarily close to the previous value P_t , which is also the upper bound of the information set of the monopolist at time $t+1$: $I_{t+1} = [\underline{v} , P_t]$.

However, the following lemma excludes that P_{t+1} be too close to P_t :

Lemma: There exists a uniform bound $k_\delta \varepsilon [x_\delta, 1]$ such that for all $x \varepsilon [0, x_\delta]$, the optimal initial price $P_0(x)$ played by the monopolist whose prior information is that $v \varepsilon [x, 1]$ verifies: $P_0(x) < k_\delta$.

The proof of this lemma is technical and can be found in the appendix. However, the intuition is simple: By choosing his initial price $P_0(x)$ arbitrarily close to 1, the monopolist loses a lot in terms of his short-run expected profits $(1-P_0)P_0$. On the other hand, he does not gain much in terms of information if it turns out that $P_0 > v$ $(\mu^0=0)$. In this case, his information set becomes $[x, P_0] = I_1$, which is almost as large as the previous $I_0 = [x, 1]$. The gain in terms of information (learning) and also in terms of the future expected profits is, however, substantial if $P_0 < v$ $(\mu^0=1)$. But this case can only occur with a very small probability when P_0 is arbitrarily close to 1.

Now the proof of (b) is immediate: Suppose that the effective sequence of prices (P_t) is eventually decreasing (e.g., for $t \geq t_0$).

Then, either the sequence of information sets $I_t = [P_{t_0}, P_{t-1}]$ is such that $\frac{P_{t_0}}{P_t} \geq x_\delta$ for t large. In this case (P_t) would become stationary at P_t, i.e., non-decreasing, a contradiction. Or we have: $\frac{P_{t_0}}{P_t} < x_\delta$ for all t, in which case the lemma applies so that, by homogeneity of our problem, we get:

$$P_t \leq k_\delta \cdot \sup I_t = k_\delta \cdot P_{t-1}\ , \text{ for all } t \geq t_0 \quad (\text{where } k_\delta < 1\ ) \quad .$$

But this is impossible, since we know that (P_t) cannot be decreasing (for $t \geq t_0$) without being uniformly bounded below by $v > 0$.

This establishes (a) and the theorem is proved. □

We conclude this section with a few remarks:

1) We have shown, part (b) of the theorem, that except when the true valuation v is zero, the sequence of prices (i.e., the optimal intertemporal pricing policy when v is the true value) is eventually non-decreasing. This property is mainly a consequence of the fact that it is never optimal (even for a non-myopic monopolist) to choose the price P_t at any period t, too close to the upper bound of the information set I_t at that time; we must have instead: $P_t \leq k_\delta \cdot \sup I_t$, where $k_\delta < 1$. The intuition is that by setting the price too close to the upper bound, the monopoly faces too high a risk of not selling anything.

2) To characterize the pricing path $P_t(v)$ further, we need to make additional assumptions. For example, we are unable to obtain the result that $P_t(v)$ is decreasing at first and eventually constant, as in Lazear (1986), without making additional assumptions on the information set $v \in [0,1]$. In fact, we can prove that if $v \in [0,1]$ the price sequence may be sometimes increasing. This follows straightforwardly from the proposition below:

<u>Proposition 4</u>: The optimal initial price P_0 set by the monopolist whose prior information is $v \in [0,1]$ satisfies the inequality:

$$P_0 < x_\delta \quad .$$

The proof can be found in the appendix.

Whenever $v \geq P_0$ the next price, P_1, will be strictly greater than P_0 (see Proposition 3(a)). Thus, prices may actually be increasing. The reason why Lazear obtains a decreasing sequence of prices is that consumers purchase only once.[3]

3) Experimentation is costly because the firm discounts the future ($\delta<1$) and as δ tends to 1, x_δ tends to 1, so that in the limit the firm only stops experimenting when it knows the exact value of v. Now, δ can be interpreted as a measure of the frequency of price offers within a given period of time. Then one may ask what prevents the firm from making an arbitrarily large number of price offers in a given time interval. Several informal arguments can be given, which explain why the interval between two price offers is not arbitrarily small. First, the flow of demand may be irregular: for example, consumers may purchase the good only on Saturdays (or for some goods, only every Christmas). Second, demand is usually stochastic. Then the firm may have to keep its price fixed for awhile in order to separate the random component from the deterministic component. More generally, whenever demand is stochastic, the firm's inference problem is harder and it takes time to learn expected revenue at any given price.

Section IV: Continuous demand functions:

The models developed in Sections II and III are special in at least one important respect: the demand function is discontinuous. Here we consider continuous demand functions and ask whether continuity is a sufficient condition to achieve complete learning of the monopoly-price in the limit. With a continuous demand function the monopolist can learn about the demand function without incurring too high experimentation costs by keeping price-variations small. This was not possible in the examples analyzed so far. Consequently, one may conjecture, first, that since experimentation costs can be arbitrarily small with a continuous demand, the firm would only want to stop experimenting when it has learned the monopoly-price. Second, since experimentation only stops when the monopoly-price is attained one may believe that the firm will eventually learn the monopoly-price.

To illustrate the first point, consider the following modification of the demand function defined in Section III: instead of having all consumers purchase one unit when $p < v$, suppose that consumers buy $q = \min\{v-mp;1\}$ units when $p < v/m$ (see Figure 4). (We assume that m is known to the firm but that the firm's prior beliefs over v are given by the uniform distribution over $[0,1]$.)

It is straightforward to show that under the modified demand curve, complete learning is the only equilibrium outcome. The difference with the discontinuous case is that the cost of local price experimentation can be made arbitrarily small (if the firm varies its price slightly above some price at which $q=1$, it loses at worst only a small fraction of demand, whereas in the discontinuous case it could lose the entire market). Since experimentation costs can be made arbitrarily small, it is always profitable for the firm to keep on experimenting

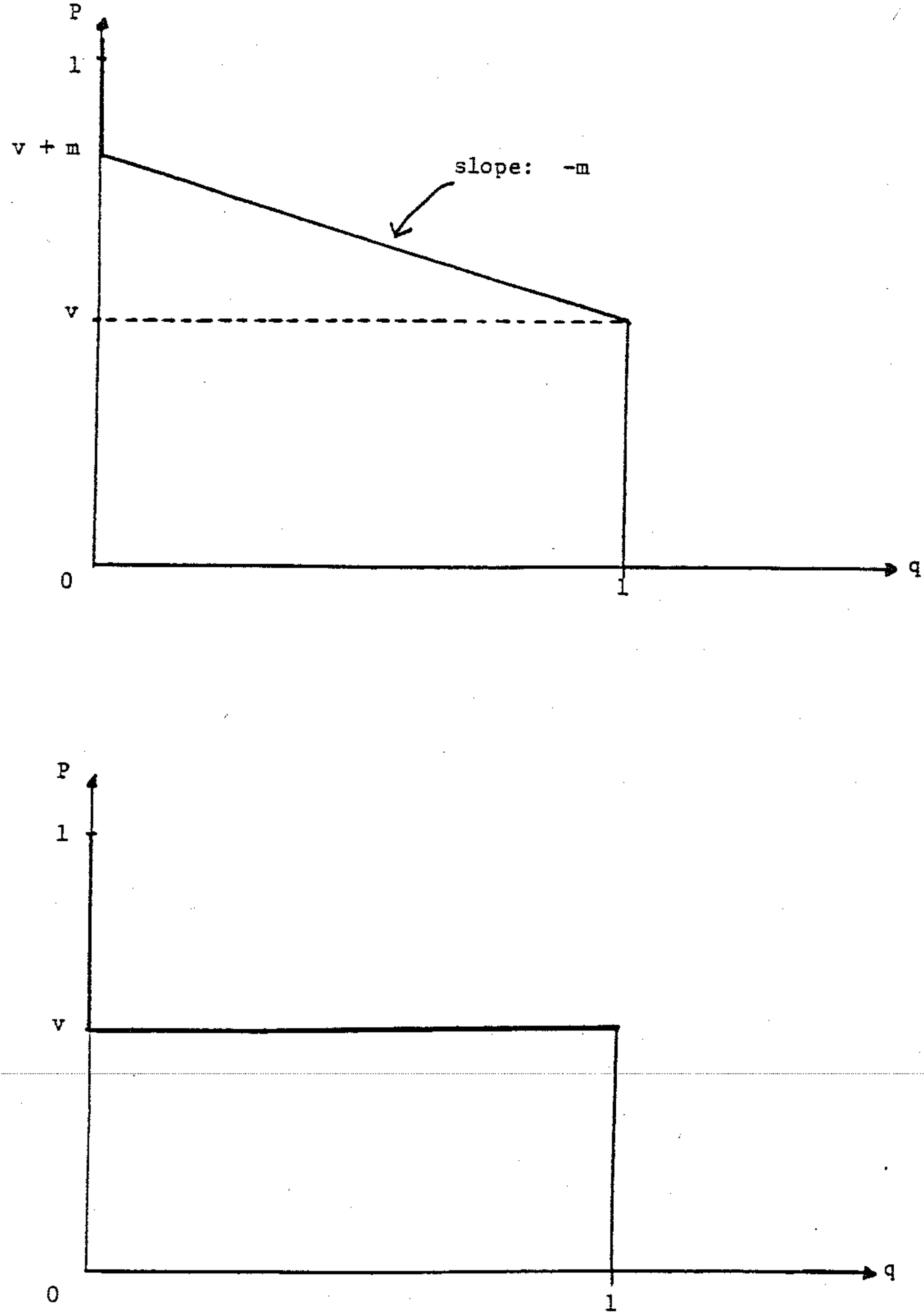

Figure 4

until it reaches a price at which the derivative of the profit-function is zero. That is, until it reaches a local optimum. It turns out that in the above example there is a unique local optimum which corresponds to the monopoly price. Thus, the firm only stops experimenting when it learns the monopoly-price.

In the above example, the key point is that the firm continues experimenting until it has reached a price at which the derivative of the profit-function is zero. For more general spaces of continuous demand functions the firm will follow the same learning rule only if <u>the profit function is known to be sufficiently smooth</u>. We establish this in the theorem below:

Let $\Pi(P,\theta)$ denote the one-shot profit function where θ is some unknown parameter. We assume that $\Pi(\cdot;\theta)$ is C^2 in P and that $\Pi(\cdot;\theta)$ and its derivative with respect to P (denoted $\Pi_1(P;\theta)$) are measurable in θ. Finally, let $\Pi_{11}(P,\theta)$ denote the second partial derivative of $\Pi(P;\theta)$ with respect to P and let H_t denote the information set of the monopolist at time $t+1$.

<u>Theorem 2</u>: Assume that $\Pi_{11}(P,\theta)$ is locally bounded in P; that is, for all P there exists $\alpha > 0$ and a neighborhood $V(P)$ such that $|\Pi_{11}(g,\theta)| < \alpha$ for all $g \in V(P)$, uniformly on $\theta \in H_t$.

Then, at each time period $t+1$ the firm either keeps on experimenting by setting $P_{t+1} \neq P_t$, or P_t is such that $E_\theta(|\Pi_1(P_t,\theta)| / H_t) = 0$.

The proof of Theorem 2 can be found in the appendix. Here we provide a sketch of the proof:

<u>Sketch of Proof</u>: The intuition of the proof is the following:

Suppose that the slope $\Pi_1(P_t,\theta)$ of the profit function Π at P_t is non-zero with positive probability given information H_t (i.e., $E_\theta(|\Pi_1(P_t,\theta)| / H_t) > 0$). Then, by experimenting with a new price $P = P_{t+1}$ close enough to P_t but different from P_t, the monopolist can obtain a very good approximation of the slope $\Pi_1(P_t,\theta)$ and in particular he can learn the sign of this slope: This, in turn, will enable him to choose a price P_{t+2} next period, such that: $P_{t+2} > P_t$ if $\Pi_1(P_t,\theta) > 0$ and $P_{t+2} < P_t$ if $\Pi_1(P_t,\theta) < 0$. The assumption that the second partial derivative $\Pi_{11}(P,\theta)$ is uniformly bounded around P_t guarantees that profits $\Pi(P_{t+2},\theta)$ cannot change too rapidly as P_{t+2} moves away from P_t; in particular the monopolist can always choose P_{t+2} so as to make sure that his short-run expected profits at this point are strictly larger than $\Pi(P_t,\theta)$: $E(\Pi(P_{t+2},\theta)|P_{t+1}=P,H_t) > \Pi(P_t, \theta)$. Now, by taking $P = P_{t+1}$ arbitrarily close to P_t, the monopolist will reduce the short-run loss on his expected profits down to zero because Π is <u>continuous</u> in P. But at the same time he will get a better estimate of the slope $\Pi_1(P_t,\theta)$ and this can only increase his expected profits at time $t+2$. Therefore, local experimentation around P_t is more profitable for the monopolist than charging the uninformative price P_t forever.

Another important feature of the demand function represented in Figure 4 is that there is only one isolated price at which the derivative of the profit-function is zero, namely the full-information monopoly-price. This feature must be preserved in more general spaces of continuous demand functions, in order to establish that experimentation only stops when the monopoly price is attained. To illustrate this we will present an example, where the profit function has a zero derivative at two isolated prices and where the firm may decide to stop experimenting even though it knows that it has not yet found the full-information monopoly-price.

In this example, the firm's profit-function, $\Pi(P)$, is assumed to take the following form:

$$\Pi(P) = \begin{cases} g(P) & \text{for} \quad P \in [0,v] \cup (v+\Delta, +\infty) \\ f(P,v) & \text{for} \quad P \in [v,v+\Delta] \quad ; \Delta > 0 \end{cases}$$

where $f(v,v) = g(v)$ and $f(v+\Delta,v) = g(v+\Delta)$ furthermore $\max_P g(P) = 1$ and the maximum is reached at $P = 1$; next $\max_P f(P,v) = 1+\lambda$ where $\lambda > 0$.

The figure below represents $\Pi(P)$:

[Figure 5 about here.]

Suppose first that the firm knows everything about its profit function. Then it is clear that the firm will choose the optimal price in $[v,v+\Delta]$ and make total intertemporal profits of $(1+\lambda) / 1-\delta$.

But when the firm is uncertain about the exact value of v we show that it may decide to always set the price $P=1$, even though it knows that this is not the full-information monopoly-price.

Let the prior distribution over v be uniform over the interval $[\underline{v},\bar{v}]$ where $1 < \underline{v} < \bar{v}$. Initially the firm can either set $P=1$ or choose some price $P \in [\underline{v},\bar{v}+\Delta]$. When the firm sets $P \in [\underline{v},\bar{v}+\Delta]$ the maximum intertemporal profits are less than or equal to

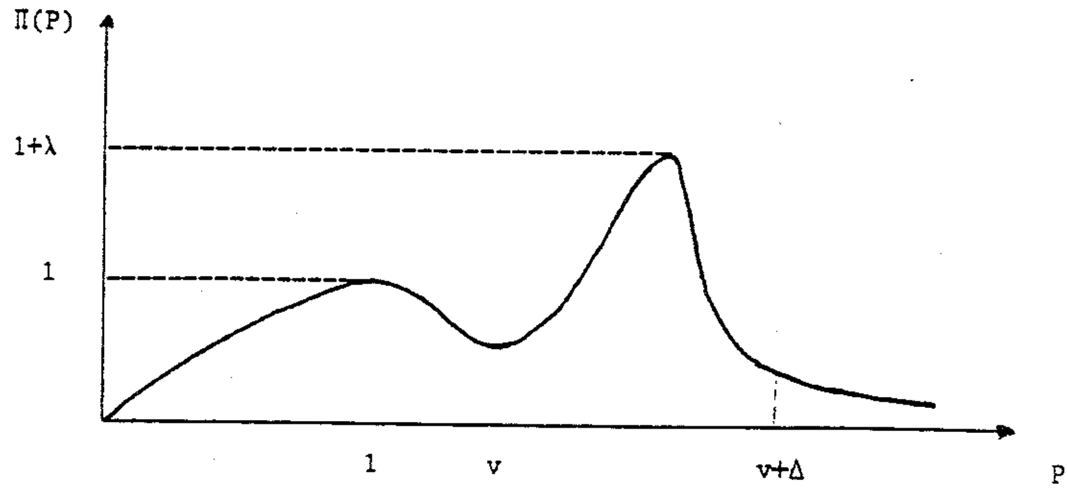

Figure 5

$$(4.1) \qquad \text{Prob}(P \,\varepsilon[v,v+\Delta])(1+\lambda) + \text{Prob}(P \not\varepsilon[v,v+\Delta])g(P) + \frac{\delta(1+\lambda)}{1-\delta}$$

The above expression is obtained by assuming first, that if $P \,\varepsilon[v,v+\Delta]$ the firm attains the maximum profits, $1+\lambda$, and second, that whatever happens in the first period the firm ends up knowing everything about its profit-function in all subsequent periods, so that its net present value of profits thereafter is given by $\frac{(1+\lambda)}{1-\delta}$.

(4.1) can be rewritten as follows:

$$(4.2) \qquad g(P) + \text{Prob}(P \,\varepsilon[v,v+\Delta])(1+\lambda-g(P)) + \frac{\delta(1+\lambda)}{1-\delta}$$

Since v is uniformly distributed on $[\underline{v},\bar{v}]$ we have $\text{Prob}(P \,\varepsilon[v,v+\Delta]) \le \frac{\Delta}{\bar{v}-\underline{v}}$, for any $P \,\varepsilon[\underline{v},\bar{v}+\Delta]$.

Thus we obtain the following upper bound:

$$(4.3) \qquad g(P)[1 - \frac{\Delta}{\bar{v}-\underline{v}}] + \frac{\Delta}{\bar{v}-\underline{v}} (1+\lambda) + \frac{\delta(1+\lambda)}{1-\delta}$$

Assuming that $g(P) \le g(\underline{v})$ for all $P \,\varepsilon[\underline{v},\bar{v}+\Delta]$ we obtain that whenever,

$$(4.4) \qquad \frac{1}{1-\delta} \ge g(\underline{v}) \; [1 - \frac{\Delta}{\bar{v}-\underline{v}}] + \frac{\Delta}{\bar{v}-\underline{v}} (1+\lambda) + \frac{\delta(1+\lambda)}{1-\delta}$$

the firm will never choose $P \in [\underline{v}, \bar{v}+\Delta]$. That is, it will stop experimenting and set the price $P=1$, even though it knows that this is not the full-information monopoly-price.

Whether or not the firm will decide to play safe and set $P=1$ forever depends on the degree of prior uncertainty about the monopoly-price (measured by the ratio $\frac{\Delta}{\bar{v}-\underline{v}}$) ; on the potential gains of experimentation (measured by λ) and finally, on the potential cost of experimentation (measured by $g(\underline{v})$) .

In the above example small price variations are not sufficient to learn about the full-information monopoly-price. Once the firm has reached the price $P=1$, it cannot learn about the optimal price without engaging in large price experimentations. Thus it cannot learn about the profit-function by incurring only arbitrarily small experimentation costs. This brings us back to the models analyzed in Sections II and III: because expected learning costs are large the firm prefers to stop experimenting before it knows all the relevant information about demand.

To summarize our discussion in this section, we have shown that continuity of the profit-function is not a sufficient condition to obtain complete learning in the limit. The profit-function must be known to be both continuous and quasi-concave. The firm will stop experimenting only when it has learned the full-information monopoly-price only if these two conditions are satisfied.

So far we have only established that if the profit-function is continuous and quasi-concave the firm will not stop experimenting before it has reached the monopoly-price. This does not imply that the learning process will eventually converge to the full-information monopoly-price. It is conceivable that although the firm never stops experimenting the price sequence generated by the learning process converges to an accumulation point which is not the full-information monopoly-price. We conjecture, however, that if the firm's learning strategy is optimal (and if the profit-function is sufficiently smooth) such an outcome can be ruled out. Intuitively, if the price sequence converges to an accumulation point which is not the monopoly-price, then the firm will eventually know that the slope of the profit-function at that point is different from zero. This information ought to induce the firm to move away from that point so that the only possible accumulation point must be the monopoly-price. We tried, without success, to provide a formal proof of convergence based on the above intuition. The difficulty in proving convergence arises from the possibility that the firm may be able to acquire most of its information about the profit-function by experimenting forever in a small neighborhood, so that the price sequence never converges to the monopoly price. We have not been able to rule out such an outcome.

Section V: Conclusion:

We have found two reasons for which it may not be in the firm's interest to continue learning until it has all the relevant information about demand: discontinuities in the demand curve and/or non-concavities in the profit-function. In either of these cases the firm cannot avoid large changes in profits while it is learning and therefore expected benefits from experimentation will be outweighted at some point during the learning process by expected costs. It is then in the firm's interest to stop learning even though it does not possess all the relevant information about demand.

We have shown that it is not necessary to assume that demand is stochastic to obtain incomplete learning results. In fact, randomness of the demand curve by itself does not imply that the firm will stop learning before it has all the relevant information about demand. Undoubtedly, though the firm's learning problem becomes worse if demand is random: First, the firm's inference problem becomes harder. Second, if there are shocks on demand, the firm does not have the same incentive to learn since the knowledge it acquires about demand will become obsolete after every shock on demand. In the extreme case where demand is independent from period to period, there is no incentive to learn anything.

We hope that the methodology developed here will be useful in studying a number of interesting extensions. Remaining in the context of a monopoly, one may ask what the consequences are of allowing consumers to adopt a strategic behavior in order to influence the firm's inferences about demand? Another interesting question is experimentation with quantity or price or a combination of both to learn most effectively about demand: When a firm experiences a shock on demand, should it first change the price and keep its sales fixed or keep the price fixed and adjust output, or use a combination of price and quantity changes? More generally, when the firm does not know the demand the question arises whether it should produce to order or determine its output first and then sell whatever it is able to sell (up to capacity). This involves a detailed study of production technologies. We plan to study two extensions in the immediate future. The first is how the stochastic nature of demand may lead to short-run price rigidity and how keeping the price fixed in the short run influences the adjustment process towards long-run equilibrium. (A firm may wish to keep its price fixed in the short run when demand is stochastic to improve its inference problems.) The second extension is learning the demand curve by duopolists.

Footnotes

1. Examples (1) and (2) are similar in spirit to Rothschild's example, which shows that when the distribution of prices across firms is unknown to consumers the optimal Bayesian search rules do not have the reservation price property (see Rothschild 1974b). His example and examples (1) and (2) above rely heavily on the property that some prices reveal a lot of information about the demand curve (or in Rothschild's model about the distribution of prices).
2. Our main results can actually be established in the more general case of a continuous density distribution $f(v)$ such that $0 < \underline{f} \leq f(v) \leq \bar{f} \leq 1$ for all $v \in [0,1]$. See Appendix.
3. An immediate consequence of Proposition 4 is that $v = 0$ will eventually be learned by the monopolist if this is the true reservation value. To be precise, $v = 0$ is such that $(P_t(0))_t$ converges to zero and most importantly, the nested sequence of information sets $I_t(v=0)$ eventually shrinks to the singleton $v = 0$. The argument goes as follows: We know that the initial price P_0 is strictly less than x_δ. If $v = 0$, the firm will not sell at price P_0 and the information set shrinks from $I_0 = [0,1]$ to $I_1 = [0,P_0]$. By homogeneity, $V_\delta(0,P_0) = P_0 V_\delta(0,1)$. Thus $P_1 = P_0^2$ and $I_2 = [0,P_0^2]$. By iteration we obtain that $P_t = (P_0^t)$ is the sequence of prices associated with $v = 0$, to which corresponds the sequence of information sets $I_t = [0,(P_0)^{t-1}]$. Since $P_0 < x_\delta < 1$, we have $(P_0^t) \to 0$, i.e., $I_t \to \{0=v\}$. □

References

Bohi, D. R. (1981), Analyzing Demand Behaviour: A Study of Energy Elasticities, The Johns Hopkins University Press, Baltimore and London.

DeGroot, M. H. (1970), Optimal Statistical Decisions, McGraw-Hill Book Co., Inc., New York.

Grossman, S., Kihlström, R., and Mirman, L. (1977), "A Bayesian Approach to the Production of Information and Learning by Doing," Review of Economic Studies 44(3): 533-548.

Lazear, E. (1986), "Retail Pricing, the Time Distribution of Transactions and Clearance Sales," American Economic Review 76(1): 14-32.

McLennan, A. (1984), "Price Dispersion and Incomplete Learning in the Long Run," Journal of Economic Dynamics and Control, (7): 331-347.

Rothschild, M. (1974), "A Two-Armed Bandit Theory of Market Pricing," Journal of Economic Theory, 9(2): 185-202.

Appendix

Proof of Theorem 2:

Let P_t be such that $E_\theta(|\Pi_1(P_t,\theta)| / H_t) > 0$, and let $V(P_t)$ denote the neighborhood of P_t where $|\Pi_{11}(P,\theta)| < \alpha$ for all $\theta \varepsilon H_t$.

Step 1: For all $P \varepsilon V(P_t)$, and for all $\theta \varepsilon H_t$:

$$\left|\Pi_1(P_t,\theta) - \frac{\Pi(P,\theta) - \Pi(P_t,\theta)}{P-P_t}\right| \leq \alpha \frac{|P-P_t|}{2} .$$

i.e., $\Pi_1(P_t,\theta)$ can be arbitrarily approximated by the slope

$$d(P) = \frac{\Pi(P,\theta) - \Pi(P_t,\theta)}{P-P_t}$$ as P moves closer to P_t.

Proof: Π being c^2, we have:

$$\Pi(P,\theta) - \Pi(P_t,\theta) = \int_{P_t}^{P} \Pi_1(q,\theta)dq$$

and:

$$\Pi_1(q,\theta) - \Pi_1(P_t,\theta) = \int_{P_t}^{q} \Pi_{11}(r,\theta)dr .$$

Now for $q \varepsilon V(P_t)$, $q > P_t$:

$$|\Pi_1(q,\theta) - \Pi_1(P_t,\theta)| \leq \int_{P_t}^{q} \alpha\, dr = \alpha(q-P_t) .$$

Therefore, if $P > P_t$, and $P \varepsilon V(P_t)$, we must have:

$$(1) \quad \Pi(P,\theta) - \Pi(P_t,\theta) \leq \int_{P_t}^{P} [\Pi_1(P_t,\theta) + \alpha(q-P_t)]dq$$

$$= \Pi_1(P_t,\theta)\cdot(P-P_t) + \alpha \cdot \frac{(P-P_t)^2}{2}$$

and similarly:

$$(2)\quad \Pi(P,\theta) - \Pi(P_t,\theta) \geq \int_{P_t}^{P} [\Pi_1(P_t,\theta) - \alpha(q-P_t)]dq$$

$$= \Pi_1(P_t,\theta)\cdot(P-P_t) - \alpha\frac{(P-P_t)^2}{2} \quad .$$

(1) and (2) suffice to establish Step 1 in the case $P > P_t$. The proof is identical when $P < P_t$. □

<u>Step 2</u>: Suppose that $E_\theta(\Pi_1(P_t,\theta)|H_t) = 0$, so that a myopic monopolist would stop experimenting at P_t . Then $\exists K > 0$ and $\mu > 0$ such that, for P close enough to P_t :

$$Pr(d(P) > K|H_t) > \mu \quad .$$

<u>Proof</u>: From Step 1, we know that for $P \in V(P_t)$:

$$(3)\quad d(P) \geq \Pi_1(P_t,\theta) - \alpha\frac{|P-P_t|}{2} \quad .$$

But by assumption, $\Pi_1(P_t,\theta)$ is non-zero with positive probability, given information H_t . In particular, $\Pi_1(P_t,\theta)$ is strictly positive with positive probability, otherwise $E_\theta(\Pi_1(P_t,\theta)|H_t)$ would be different from zero, which we assumed away in this Step 2. This means that we can find $L > 0$ and $\mu > 0$ such that:

$$(4)\quad Pr(\Pi_1(P_t,\theta) > L|H_t) > \mu \quad .$$

From (3) and (4) it turns out that we can always choose P close enough to P_t .

$$Pr(d(P) > L/2|H_t) > \mu \quad .$$

It suffices now to take $K = L/2$ in order to complete the proof. □

<u>Step 3</u>: Under the assumption of Step 2, for $P = P_{t+1}$ close enough to P_t but $P \neq P_t$, there exist $F > 0$ and a new price $P_{t+2} = P_{t+2}(P) \in V(P_t)$ such that:

$$E_\theta(\Pi(P_{t+2},\theta)|H_t \text{ and } P) \geq \Pi(P_t,\theta) + F \quad .$$

<u>Proof</u>: Let ε be a positive number smaller than K : $0 < \varepsilon < K$.

For $P = P_{t+1}$ close enough to P_t, $\alpha \frac{|P-P_t|}{2}$ can be made smaller than ε, so that by Step 1:

$$\Pi_1(P_t,\theta) > d(P) - \varepsilon \quad .$$

Now, for any $q \in V(P_t)$, we have:

$$\Pi(q,\theta) \geq \Pi(P_t,\theta) + (d(P)-\varepsilon)(q-P_t) - \alpha \frac{(q-P_t)^2}{2} \quad .$$

Let $P_{t+2} \in V(P_t)$ be defined as follows:

$$\begin{cases} \cdot \; P_{t+2} = P_t + \dfrac{d(P)-\varepsilon}{\alpha} & \text{if } 0 < \dfrac{d(P)-\varepsilon}{\alpha} < r \\ & \text{where } r = \sup\limits_{q \in V(P_t)} (q-P_t) \quad . \\ \cdot \; P_{t+2} = P_t + r & \text{if } \dfrac{d(P)-\varepsilon}{\alpha} > r \quad . \\ \cdot \; P_{t+2} = P_t & \text{otherwise.} \end{cases}$$

We have:

$$\Pi(P_{t+2},\theta) \geq \Pi(P_t,\theta) + \frac{\alpha\tau^2}{2} \quad \text{if} \quad d(p)-\varepsilon > 0$$

$$\text{where} \quad \tau = \min(r, \frac{d(P)-\varepsilon}{\alpha}) \quad .$$

and $\Pi(P_{t+2},\theta) \geq \Pi(P_t,\theta)$ if $d(P) < \varepsilon$.

$$\Rightarrow \quad E(\Pi(P_{t+2},\theta)|H_t \text{ and } P) \geq Pr(d(P)>\varepsilon)\cdot[\Pi(P_t,\theta) + \frac{\alpha\cdot\tau^2}{2}]$$

$$+ Pr(d(P)\leq\varepsilon)\cdot\Pi(P_t,\theta)$$

$$\geq \Pi(P_t,\theta) + Pr(d(P)>\varepsilon)\cdot \frac{\alpha\cdot\tau^2}{2} \quad .$$

We know from Step 2 that:

$$Pr(d(P)>K) > \mu \quad .$$

Given that $\varepsilon < K$, we necessarily have: $Pr(d(P)>\varepsilon) > \mu$.

This establishes Step 3, with $F = \mu \frac{\alpha\tau^2}{2} > 0$. □

Now the proof of Theorem 2 goes as follows:

- Either $E_\theta(\Pi_1(P_t,\theta)|H_t)$ is different from zero, in which case even a myopic monopolist would keep on experimenting at time $t+1$ (a fortiori, would the non-myopic monopolist do so),
- or $E_\theta(\Pi_1(P_t,\theta)|H_t) = 0$; in which case, by experimenting through $P = P_{t+1}$ different from P_t but close enough to P_t, the monopolist can obtain an expected intertemporal profit $\Pi_\delta(P|H_t)$, where, from Step 3:

$$\Pi_\delta(P|H_t) \geq \Pi(P,\theta) + \frac{\delta}{1-\delta}(\Pi(P_t,\theta) + F) \quad .$$

Now, by continuity of Π, we have: $\lim_{P_t \to P} \Pi(P,\theta) = \Pi(P_t,\theta)$, so that:

$$\lim_{P \to P_t} \Pi_\delta(P|H_t) \geq \frac{\Pi(P_t,\theta)}{1-\delta} + \frac{\delta F}{1-\delta}$$

$$= \Pi_\delta(P_t|H_t) + \frac{\delta F}{1-\delta} \quad .$$

Therefore, given that $\delta > 0$, the monopolist is better off experimenting locally around P_t than charging this uninformative price forever.

This concludes the proof. □

<u>Proof of Lemma</u>:

<u>There exists a uniform bound $k_\delta \in [x_\delta,1]$ such that for all $x \in [0,x_\delta]$</u>, the optimal initial price $p_o(x)$ played by the monopolist whose prior information is that $v \in [x,1]$ verifies:

$$p_o(x) \leq k_\delta \quad .$$

<u>Proof</u>: Let $x \in [0,x_\delta]$ and suppose that the optimal initial price $p_o(x) \in [x_\delta,1]$. Then we know that $V(p_o,1) = p_o/1-\delta$ so that:

$$(1-x)\ V(x,1) = \frac{(1-p_o)p_o}{1-\delta} + \delta(p_o - x)V(x,p_o)$$

and, using $V(x,p_o) < V(x,1)$,

$$V(x,1) < \frac{(1-p_o)p_o}{(1-\delta)(1-\delta p_o-(1-\delta)x)} .$$

Since $x < x_\delta$ and $V(x,1) > V(o,1) \geqq \frac{1}{(4-\delta)(1-\delta)}$ (Proposition 3), the price p_o must verify:

$$(1-p_o)\ p_o > \frac{1}{4-\delta}\ [\ \frac{1}{2-\delta} - \delta p_o\ ] \quad \text{or}$$

$$0 > p_o^2 - \frac{\delta}{4-\delta}\ p_o + \frac{1}{(4-\delta)(2-\delta)} = g(p_o)$$

. It is straightforward to verify that for all $\delta < 1$, $g(x) < 0$ and $g(1) > 0$. Thus, there exists $1 > k_\delta > x_\delta$, such that:

$$p_o \in [k_\delta, 1] \Rightarrow g(p_o) > 0 .$$

Therefore, if the optimal price p_o is larger than x_δ , it must be strictly less than k_δ .

The lemma is proved. □

Proof of Proposition 4:

First step: The optimal initial price p_o verifies:

$$p_o \leqq x_\delta = \frac{1}{2-\delta} .$$

Proof: Suppose that $p_o > \frac{1}{2-\delta}$. We have

$$(1):\ V_\delta(0,1) = \max_{P\epsilon[0,1]} \{P(1-P) + \delta(1-P)V_\delta(P,1) + \delta P\ V_\delta(0,P)\} .$$

The f.o.c. corresponding to this maximization are given by:

$$1 - 2P_0 - \frac{\delta P_0}{1-\delta} + \frac{\delta(1-P_0)}{1-\delta} + 2\delta P_0 V_\delta(0,1) = 0$$

(we use the fact that $V_\delta(P_0,1) = \frac{P_0}{1-\delta}$, since $P_0 > \frac{1}{2-\delta}$, by hypothesis).

The above f.o.c. are rewritten as follows:

$$f(P_0) = \frac{1-2P_o}{1-\delta} + 2\delta P_0 V_\delta(0,1) = 0 \quad \text{, so} \quad f(P_0) = 0 \quad \text{, for} \quad P_0 > \frac{1}{2-\delta} \text{, if}$$

and only if $V_\delta(0,1) > \frac{1}{2(1-\delta)}$. But, since $P_0 > \frac{1}{2-\delta}$, we know that

$$V_\delta(0,1) = P_0(1-P_0) + \delta(1-P_0)\frac{P_0}{1-\delta} + \delta P_0^2\, V_\delta(0,1)$$

$$\Rightarrow V_\delta(0,1) = \frac{P_0(1-P_0)}{(1-\delta P_0^2)(1-\delta)} = g(P_0) \quad .$$

It is straightforward to compute that for $P_0 > \frac{1}{1+\sqrt{1-\delta}}$, $g'(P_0) < 0$.

Therefore, since $x_\delta = \frac{1}{2-\delta} < \frac{1}{1+\sqrt{1-\delta}}$, we have:

$g(P_0) < g(x_\delta) = \frac{1}{(4-\delta)(1-\delta)}$. That is, $V_\delta(0,1) < \frac{1}{(4-\delta)(1-\delta)} < \frac{1}{2(1-\delta)}$,

a contradiction.

Step 1 is proved. □

<u>Second step</u>: $P_0 < x_\delta$.

<u>Proof</u>: We know, from Step 1 above, that $P_0 \le x_\delta$. Now suppose that $P_0 = x_\delta$. We would have:

$$W_\delta(0) = \left(1 - \frac{1}{2-\delta}\right)\frac{1}{2-\delta} + \delta\left(1 - \frac{1}{2-\delta}\right)\frac{1}{(2-\delta)(1-\delta)} + \frac{\delta}{(2-\delta)^2}\cdot W_\delta(0)$$

$$> W_\delta(0) = \frac{1}{(2-\delta)^2-\delta} = \frac{1}{(4-\delta)(1-\delta)} = M_\delta(0,1) \quad \text{, i.e., the}$$

<u>myopic</u> intertemporal profit; but this is impossible by Proposition 3. This completes the proof. □