working paper
department
of economics

Learning Through Price Experimentation
By a Monopolist Facing Unknown Demand

By
Philippe Aghion,
Patrick Bolton,
and
Bruno Jullien

Number 491

March 1988

massachusetts
institute of
technology
50 memorial drive
cambridge, mass. 02139

Learning Through Price Experimentation By a Monopolist Facing Unknown Demand

By
Philippe Aghion,
Patrick Bolton,
and
Bruno Jullien

Number 491 March 1988

August 1987, revised March 1988

LEARNING THROUGH PRICE EXPERIMENTATION BY A MONOPOLIST FACING UNKNOWN DEMAND

by

Philippe Aghion**

Patrick Bolton***

and

Bruno Jullien***

* We are greatly indebted to Jerry Green, Andreu Mas-Colell, Eric Maskin, Iraj Saniee and Jean Tirole. We have benefited from many helpful comments and discussions with seminar participants at Stanford, Harvard, Chicago, UCLA, Cornell and Berkeley. Philippe Aghion and Bruno Jullien gratefully acknowledge financial support from the Sloan Foundation Dissertation Fellowship.

** Massachusetts Institute of Technology.

*** Harvard University.

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 the 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

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

Introduction

In almost all the 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 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 prices and output 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. residencial electricity demand (see Bohi [1981] for a survey) at best provide information about price-elasticity 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 its prices at every period so as to maximize profits. The main issues here are: how does learning considerations affect the firm's pricing decisions and assuming that demand remains stable over time when will the firm stop experimenting.

The firm's learning process through time can be viewed as an <u>adjustment</u>

process toward 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 toward the new equilibrium by experimenting with prices.[1] 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 toward understanding learning by experimentation, we consider the case of a monopoly. Some of the issues arising with oligopoly are addressed in Aghion-Espinoza- Jullien [1988]. 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 firm 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.[2] They are, however, able to use data about past sales to update their priors. We analyze both types of behavior.

The existing literature on this question comprises only a handful of papers: Rothschild [1984]; Grossman- Kihlström- Mirman [1977]; McLennan [1984]; Lazear [1986] and Alpern-Snower [1987a,b]. Most of these studies consider examples with unknown stochastic demand functions. As a result the analysis becomes quite involved and sometimes the underlying principles driving some of the results are difficult to isolate. It turns out, however, that some of the most important conclusions reached in this literature can be easily derived in the following example with deterministic demand.

A monopolist produces a non-storable good at zero unit cost and sells in a market composed of two types of consumers: those who attach a high reservation value to the good and those who attach a low value. Each consumer purchases at most one unit of the good every period. The reservation value of each type of consumer are given by $v_1 > v_2 > 0$, and the proportion of high-value consumers is $\mu \in [0,1]$. When v_1, v_2 and μ are known to the firm, but the firm cannot identify the type of the buyer, then the monopoly-price set at each period is given by:

$$\begin{cases} P_t = v_1 \text{ if } \mu . v_1 \geq v_2 \\ P_t = v_2 \text{ if } \mu . v_1 < v_2 \end{cases}$$

Suppose now that μ is initially unknown to the firm and that the prior distribution of μ is the uniform distribution on [0,1]. Our example has the simple feature that if the firm sets the initial price $P_1 = v_1$ it learns μ immediatly, but if it sets $P_1 = v_2$ it does not learn anything about μ. The firm's inference problem is as simple as it can be. It is easy to show that even in this simple example the optimal policy for the firm may be not to experiment to learn μ (i.e., to set the price v_2 in each period). Let δ be the firm's discount factor and assume that the firm can sell for an infinite number of periods. A straightforward calculus shows that the optimal pricing policy for the firm is to set the price $P_t = v_2$ for all t whenever:

(2) $v_1 \leq v_2 \ (1+\sqrt{1-\delta}\ )$.[3]

It may well be, however, that the true proportion of high-value consumers μ is strictly greater than v_2/v_1 , so that the full information monopoly-pricing policy would be to set $P_t=v_1$ for all t. The true proportion will never be learned by the firm whenever (2) is satisfied so that the monopolist may set the wrong price forever. This is essentially the main result obtained by Rothschild, although he worked with stochastic demands and used two-armed bandit theory (see Degroot [1970]) to prove his result.

Our example also yields the conclusion that a firm that takes into account the value of information obtained from experimentation tends to set higher prices initially than a myopic firm, who only takes into account current expected profits in determining its pricing decision. The myopic firm chooses:

$$(3) \begin{cases} P_1=v_2 \text{ if } & v_1 \leq 2.v_2 \\ P_1=v_1 \text{ otherwhise} \end{cases}$$

while the far-sighted firm chooses $P_1=v_2$ only if (2) is satisfied. Such a result has been first established in a different model by Grossman-Kihlström-Mirman.[4]

Our model clearly illustrates that Rothschild's result on incomplete learning depends on the assumption of a discrete price-grid. In other words it is important to assume that the firm is not allowed to experiment through small price-variations or that it cannot learn anything if it does not move its price much. McLennan provides an interesting example where the firm can experiment with small price variations. He shows, however, that this is not the most efficient way of learning. It is best for the firm to engage in price-experimentation with sufficiently large price-swings in order to learn the most about demand. As a result, the costs of price experimentation are bounded away from zero and McLennan shows that when the anticipated value of information obtained from experimentation is small the firm prefers not to experiment.

McLennan's result relies crucially on the stochastic nature of demand. One is thus led to believe (incorrectly) that incomplete learning can only arise when demand is noisy.

One of the main objectives of our paper is to show that incomplete learning results do not necessarily depend on assumptions about the discreteness of the price-grid or the stochastic nature of demand. Thus, we present an example in section I where the firm can experiment and learn about a deterministic demand with arbitrarily small changes in prices, in which the firm will almost surely stop experimenting before it knows all the relevant information about demand. Section II then addresses the general question of what structure of deterministic demand functions leads to results of incomplete learning. This is a first attempt at setting out a general theory about price-experimentation by monopolists. Most of the existing literature only looks at specific examples of demand functions. This section attempts to go further by looking at a general class of deterministic demand functions.

Another objective of our paper is to convince the reader of the usefulness of our simplifying assumption that the firm faces a deterministic demand. This greatly increases our ability to analyze problems of learning by experimentation. Thus, in the example analyzed in section I we can get an almost complete analytic solution. We can characterise the learning process and study the adjustment process of the firm overtime. Alpern-Snower and Reyniers [1987a,b,c] have adopted the same approach and have thus been able to find exact solutions to the firm's optimal adjustment process in different models.

I. An example of learning through price experimentation with deterministic demand.

We shall consider a market where all consumers purchase at most one unit each period and have the same reservation value, v. The firm does not know v initially and given any price it sets, it will either serve the entire market and learn that v is greater than its price, or it will not serve any customer and learn that v is less than its price.

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 that 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]. Consumers purchase at most one unit every period and do not behave strategically (i.e., they buy whenever $P_t \leq v$).[5] The firm produces the good at a 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].[6]

We begin with the decision problem faced by a non-myopic monopolist: for any pair (x,y) where $0 \leq x < y \leq 1$, we denote $V_\delta(x,y)$ 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 be described as follows: at time t=0, it chooses some price $P_0 \in [0,1] = I_0$. 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 $\mu^0=0$). If $\mu^0=1$, the monopolist can update his information about v to $v \in [0,P_0] = I_1$. Similarly, if $\mu^0=0$, the monopolist's information becomes $v \in [P_0,1] = I_1$. Now, the optimal choice of the initial price P_0 made by a non-myopic monopolist is the solution to the following maximization program:

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

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:

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

As in most dynamic programming problems, it is useful to define 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.

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

$$\text{(B)}\quad V(x,y) = \max_{z\in[x,y]} \frac{1}{y-x}\{(y-z)z + \delta.(y-z)V(z,y) + \delta.(z-x)V(x,z)\}$$

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

$$\pi(z,x,y) = z.(\frac{y-z}{y-x}) + \delta(\frac{y-z}{y-x})V_\delta(z,y) + \delta(\frac{z-x}{y-x})V_\delta(x,z)$$

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 to 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: $\Sigma\delta^t.\pi_t(x,y)$, where $0 \le \pi_t(x,y) \le 1$ 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: $v \in [x,y]$.)

Let B_0 denote the space of bounded functions on $[0,1]^2$. B_0 is a Banach space for the uniform norm $|.|_\infty$ defined by:

$$|f|_\infty = \sup_{x \in [0,1]^2} |f(x)| < \infty \text{ for all } f \in B_0$$

Next, consider the following mapping Π:

$$\Pi : B_0 \longrightarrow B_0$$

$$V \longrightarrow \Pi(V)$$

where:

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

$$= \max g_V(x,y,z) \quad .$$

We know that $V_\delta \in B_0$ 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' belong to B_0, and

$$z_V(x,y) = \arg\max_{z \in [x,y]} \; g_V(x,y,z).$$

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:

$$|\Pi(V)(x,y)-\Pi(V')(x,y)| \leq \delta\Big(\frac{z_V(x,y)-x}{y-x}\,[V(x,z_V(x,y))-V'(x,z_V(x,y))] + \frac{y-z_V(x,y)}{y-x}\,[V(z_V(x,y),y)-V'(z_V(x,y),y)]\Big)$$
$$\leq \delta|V-V'|_\infty.\Big[\frac{z_V(x,y)-x+y-z_V(x,y)}{y-x}\Big]$$
$$\leq \delta.|V-V'|_\infty .$$

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 complete the proof of Proposition 1.

QED

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 $|.|_\infty$ 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 proove this corollary, to show that for all $V \in B_0$ where V is homogeneous of degree 1 and convex, that $\Pi(V)$ inherits the same properties. So, let V verify these properties:

(a) Π(V) is homogeneous of degree one.

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

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

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

(b) Π(V) is convex.

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

$$\Pi(V)(x,y) = \max_q \{h(q,x,y) = (1-q).((y-x)q+x) + \delta(1-q)V((y-x)q+x,y) + \delta q.V(x,(y-x)q+x) \}.$$

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

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

(b) for each (x,y), $h(q,x,y)$ is continuous in q.

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

The fact that V_δ is increasing is elementary.

QED

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 so as to maximize:

$$\Pi(P_1, I_1) = Pr(\mu^1 = 1 \; / \; P_1 \text{ and } v \in I_1).P_1$$
$$= Pr(v \geq P_1 \; / \; v \in I_1 = [1/2, 1]).P_1$$
$$= 2(1 - P_1)P_1 \; .$$

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 set $P_t = 1/2$ for all $t \geq 1$. In other word, it stops experimenting from period one on.

What happens 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 to 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) $V_\delta(0,1) > M_\delta(0,1)$, where

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

Proof: Let $W_\delta(x) = V_\delta(x,1)$ for all $x \in [0,1]$. (By homogeneity of degree one we have: $V_\delta(x,y) = yW_\delta(x/y)$ 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 $y \in [x,1]$. Then, from the Bellman equation (B), we must have:

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

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

$$1 - 2y + \delta z = 0 \text{ , i.e., } y = \frac{1+\delta z}{2} .$$

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

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

$$y^* = \arg\max G(y) = z \text{ , i.e., } W_\delta(z) = G(z) = \frac{z}{1-\delta} .$$

This proves that $V(x,1) = x/1-\delta$ is the solution of the Bellman equation (B) on the set $\{x\in[0,1]/x\geq x_\delta\}$, 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 1 that V_δ is the unique solution of this Bellman equation on [0,1], we necessarily have:

$$V_\delta(x,1) = \frac{x}{1-\delta} \text{ for all } x \in [x_\delta,1], \text{ and}$$
$$V_\delta(x,1) \neq \frac{x}{1-\delta} \text{ for } x < x_\delta \text{ , x close to } x_\delta .$$

Insert Figure 2 about here.

But since V_δ is <u>convex</u> (corollary), 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 with 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 firm sets $P_t = 1/2$ forever, i.e.:

$$M_\delta(1/2,1) = \frac{1/2}{1-\delta} .$$

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:

$$M_\delta(0,1/2) = \frac{1}{2} M_\delta(0,1) \ .$$

We have:

$$M_\delta(0,1) = \frac{1}{2} \left(\frac{1}{2} + \delta M_\delta(\tfrac{1}{2},1)\right) + \frac{1}{2} \delta M_\delta(0,\tfrac{1}{2})$$
$$= \frac{1}{4} + \frac{\delta}{4(1-\delta)} + \frac{1}{4} \delta M_\delta(0,1)$$
$$\Rightarrow \qquad M_\delta(0,1) = \frac{1}{(4-\delta)(1-\delta)}.$$

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

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

where the RHS represents the non-myopic firm's profit when $P_0=1/2$. By homogeneity we have:

$$V_\delta(0, \tfrac{1}{2}) = \frac{1}{2} V_\delta(0,1).$$

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

$$V_\delta(\tfrac{1}{2} , 1) > \frac{1/2}{1-\delta} \ .$$

Therefore:

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

i.e., $V_\delta(0,1) > \dfrac{1}{(4-\delta)(1-\delta)} = M_\delta(0,1)$. QED

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 sure that it is setting the best possible price. Let us denote $(P_t(v))$ the sequence of prices played by the monopolist when the true reservation value is v and its prior information is $v \in [0,1]$.

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 is zero, the nested sequence of information sets $I_t = [\underline{v}_t, \bar{v}_t]$ will never converge to the singleton $\{v\}$. The lenght $\ell(I_t)$ will remain bounded away from zero as $t \to +\infty$.

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

Insert Figure 3 about here.

Although the sequence (P_t) may sometime 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 2(a), we know that the monopolist will set $P_t = \underline{v}_t$ for all $t \geq T$. This implies that $I_t = I_T$ for all $t \geq T$, a contradiction to 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 decreasing and bounded below by v (if $P_t < v$ for some t, then necessarily $P_{t+1} \geq P_t$ and P_t would not be decreasing!). Without loss of generality, we can assume that the whole sequence (P_t) is decreasing for t large enough. The sequence 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 arbitrarily close to the previous price 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 exclude that P_{t+1} be too close to P_t:

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

The proof of this lemma is technical and can be found in the appendix. However, the intuition is simple: By choosing the initial price $P_0(x)$ arbitrarily close to 1, the monopolist loses a lot in terms of his short-run expexted profits $P_0(1-P_0)/(1-x)$. On the other hand, he does not gain much in terms of information if it turns out that $P_0 > v$ (μ^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 $P_0 < v$ (μ^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 (a) is immediate: Suppose that the effective sequence of prices (P_t) is eventually decreasing (e.g., for t≥T). Then, either the sequence of information sets $I_t=[\underline{v}_T,P_t]$ is such that $\underline{v}_T/P_t \geq x_\delta$ for some t. In this case (P_t) would become stationary at $\underline{v}_T$, i.e., non-decreasing, a contadiction. Or we have $\underline{v}_T/P_t < x_\delta$ for all t, in which case the lemma applies so that, by homogeneity of our problem, we get:

$P_t\leq k_\delta.\sup I_t = k_\delta.P_{t-1}$, for all $t \geq T$.

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

QED

We conclude this section with a few <u>remarks</u>.

1) We have shown, part (a) of the theorem, that except when the true

value 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: $P_t \leq k_\delta . \sup I_t$, where $k_\delta < 1$. The intuition is that by setting the price too close to the upper bound, the monopolist faces too high a risk of not selling anything.

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

Proposition 3: The optimal initial price P_0 set by the monopolist whose prior information is $v \in [0,1]$ satisfies the inequality:

$$P_0 < x_\delta .$$

The proof can be found in the appendix.

Whenever $v \geq P_0$ the next price, P_1, will be stricly greater than P_0 (see Proposition 2(a)). Thus, prices may actually be increasing. The reason why Lazear obtains a decreasing sequence of prices is that consumers purchase only once.

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 experimemting 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 a while 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.

II. Continuous demand functions:

The model developed in Section I is special in at leat 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 I: instead of having all consumers purchase one unit when $p < v$, suppose that consumers buy $q = \min\{1, 1+\frac{v-p}{m}\}$ 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].)

Insert Figure 4 about here.

It is straightforward to show that under the modified demand curve, complete learning is the only equilibrium outcome. The difference with the discontinuous demand case is that the cost of local 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). Therefore, it is always profitable for the firm to keep on experimenting 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 the profit function is concave so that the firm only stops experimenting when it learns the monopoly-price.

In the above example, the key point is that the firm continues experimenting unless it is sure that the derivative of the profit function is zero. For more general spaces of continuous demand functions the firm will follow the same learning rule provided <u>the profit function is known to be sufficiently smooth</u>. This is established in the theorem below:

Let $\Pi(P,\theta)$ denote the one-shot profit function where θ is some unknown parameter. We assume that $\Pi(.,\theta)$ is C^2 in P and that $\Pi(.,\theta)$ 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 derivative of $\Pi(P,\theta)$ with respect to P and let H_t denote the information set of the monopolist at time t+1.

Theorem 2: Assume that $\Pi_{11}(P,\theta)$ is locally uniformly bounded for all 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 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:

Sketch of the proof: 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 <u>slope</u> $\Pi_1(P_t,\theta)$ 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 profit $\Pi(P_{t+2},\theta)$ 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 profit at time t+2. Therefore, local experimentation around P_t is more profitable for the monopolist than charging the uninformative price P_t forever.

An 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 point 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:

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

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

Figure 5 below represents $\Pi(P)$:

Insert 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:

$$(4.1) \qquad \text{Prob}(P\in[v,v+\Delta])(1+\lambda) + \text{Prob}(P\notin[v,v+\Delta])g(P) + \frac{\delta(1+\lambda)}{1-\delta}$$

The above expression is obtained by assuming first that if $P \in [v,v+\Delta]$ 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.

Since v is uniformly distributed on $[\underline{v},\bar{v}]$, we have:

$$\text{Prob}(P\in[v,v+\Delta])\leq \Delta/(\bar{v}-\underline{v}).$$

Thus we obtain the following upper bound:

$$(4.2) \quad g(p).[1-\Delta/(\bar{v}-\underline{v})] + \Delta.(1+\lambda)/(\bar{v}-\underline{v}) + \delta.(1+\lambda)/(1-\delta)$$

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

$$(4.3) \quad \frac{1}{1-\delta} \geq g(\underline{v}).[1-\Delta/(\bar{v}-\underline{v})] + \Delta.(1+\lambda)/(\bar{v}-\underline{v}) + \delta.(1+\lambda)/(1-\delta)$$

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 $\Delta/(\bar{v}-\underline{v})$); 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 example developed in the introduction: 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. <u>The profit-function must be known to be both continuous and quasi-concave</u>. The firm will stop experimenting when it has learned the full-information monopoly-price only if these two conditions are satisfied.

So far, all we have established is that provided 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.

III. 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 both cases the firm cannot avoid incurring experimentation costs that may be larger than the expected benefits from experimentation. It is then in the firm's interest to stop learning even though it does not possess all the relevant information about demand. Under those circumstances, the firm will generally end up setting a price different from the full-information monopoly-price. Such incomplete learning results are important since they suggest that existing theories of monopoly pricing cannot be viewed as representations of a firm's pricing behavior once this firm's learning process has settled down. 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, whenever learning is incomplete in equilibrium, we have a situation where the long-run equilibrium is history-dependent.

Another conclusion of our analysis is 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 inference problem becomes harder if demand is random: firstly, the firm may have to set the same price for many periods before it can infer the value of expected profits at this price with sufficient accuracy, using the law of large numbers. The stochastic nature of demand increases the experimentation time necessary to acquire any given amount of information. Secondly, what makes the learning problem particularly difficult to analyze when the firm faces a stochastic demand is the following observation: under a deterministic demand the firm's choice problem at any period is one of deciding between staying at an uninformative price forever or experimenting (and then pick the optimal experimentation price); when the firm faces a stochastic demand there may not always be such an uninformative price. In particular, by setting the same price forever, the firm may still be able to learn about demand over time. As a result, the firm's problem is no longer one of deciding between stopping to experiment or continuing to experiment. (For a detailed discussion of the stochastic case, see Jullien [1988].)

We hope that the methodology developed in this paper 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 both quantities and prices. When a firm experiences a shock on demand, should it change the price and keeps its volume of sales fixed, or else keep the price fixed and adjust output, or instead 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.

Another natural extension is to investigate the same type of learning problem in an oligopolistic context, where several competing firms face uncertainty about demand. A first step in that direction is taken by Aghion-Espinosa-Jullien [1988], which analyzes how learning considerations affect the equilibrium prices in a dynamic duopoly game.

References

Aghion,P.,Bolton,P., and Jullien,B. [1987], "Learning Through Price Experimentation by a Monopolist Facing Unknown Demand", UC Berkeley Working Paper 8748.

Aghion,P.,Espinosa,M., and Jullien,B. [1988], "Dynamic Duopoly Games with Learning Through Market Experimentation", Mimeo, Harvard University.

Alpern,S., and Snower,D.J. [1987a], "Inventories as an Information Gathering Device", ICERD Discussion Paper 87/151, London School of Economics, London.

Alpern,S., and Snower,D.J. [1987b], "Production Decisions under Demand Uncertainty: the High-Low Search Approach", ICERD Discussion Paper 223, London School of Economics, London.

Bohi,D.R. [1981], Analyzing Demand Behaviour: A Study of Energy Elasticities, The John 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.

Jullien,B. [1988], "Three Essays on the Theory of Choice Under Uncertainty and Strategic Interaction," Ph.D. Dissertation, Harvard University.

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.

Reyniers,D.J. [1987a], "A High-Low search Algorithm for a Newsboy Problem with Delayed Information Feedback," Mimeo, London School of Economics.

Reyniers,D.J. [1987b], "Interactive High-Low Search: the Case of Lost Sales." Mimeo, London School of Economics.

Reyniers,D.J. [1987c], "A Sequential Search Version of the Newsboy Problem," Mimeo, London School of Economics.

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

Footnotes

[1]Alternatively it could experiment with quantities. Quantity experimentation is investigated by Alpern-Snower [1987a,b]. In their model when the firm overproduces it learns exactly by how much it has overproduced, whereas if it underproduces it does not know exactly by how much it has underproduced. This type of information acquisition is a case of what Alpern and Snower call "high-low" search. (For other applications of the "high-low" search approach, see Reyniers [1987a,b,c].)

[2]A myopic firm can be viewed as a simple representation of a sequence of short-lived firms, each one having access to the data about past firms' pricing decisions and sales.

[3]The expected net present value of profits when the firm sets $P_1=v_1$ is:

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

When it sets $P_1=v_2$ it is simply: $\Pi_2 = \frac{v_2}{1-\delta}$.

[4] It can be shown, however, that the Grossman-Kihlström-Mirman result is not robust. Consider the following generalization of our example: There are now four types of buyers with reservations values $v_1>v_2>v_3>v_4$ and the proportions of each type are given by μ_1,μ_2,μ_3,μ_4. There are two possible demand functions that the firm can face: In state 1 the proportions are given by $(\mu_1,\mu_2,\mu_3,\mu_4) = (2/3,0,1/3,0)$, in state 2 they are given by $(0,2/3,0,1/3)$. Let α be the prior probability of state 1 and assume that $2v_2/3 > \alpha v_3+(1-\alpha)2v_3/3$ and that $v_2 > \alpha v_1$. Then one can show that for a given small α, if the discount factor is large enough the non-myopic firm will set a lower initial price than a myopic firm. For details, see Aghion-Bolton-Jullien [1987].

[5] Lazear [1986] looks at basically the same model, except that consumers are assumed to purchase the good only once. If the price is too high today they may decide to wait until tomorrow before purchasing. Once they buy, however, they drop out of the market.

[6] Our main results can actually be established in the more general case of a continuous density distribution $f(v)$. See 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 $\theta \in H_t$.

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

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

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

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

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 $q \in V(P_t)$, $q > P_t$:

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

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

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

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

and similarly:

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

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

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

Step 2: 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 :

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

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

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

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

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. □

Step 3: 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:

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

Proof: 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. □

Proof of Lemma:

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

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

Proof: 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 $V(x,1) > V(o,1) \geq \frac{1}{(4-\delta)(1-\delta)}$ (Proposition 3), the price p_o must verify:

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

$$0 > p_o^2 - \frac{\delta}{4-\delta}\ p_o + \frac{1}{(4-\delta)(2-\delta)} = g(p_o)$$. 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 3 :

First step: The optimal initial price p_o verifies:

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

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

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

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:

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

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. □

Second step: $P_0 < x_\delta$.

Proof: We know, from Step 1 above, that $P_0 \leq x_\delta$. Now suppose that $P_0 = x_\delta$. We would have:

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

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

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

Extending Model II to the case where the reservation value v is distributed according to a general density $f(v)$ on $[0,1]$.

Here we generalize our treatment of Model II by assuming that the prior information of the monopolist on the reservation value v at time $t=1$ is that v is distributed on $[0,1]$ according to a density f, not necessarily uniform, but such that

$$0 < \underline{f} \leq f(v) \leq \bar{f} \quad , \text{ for all } v \in [0,1] \quad .$$

(We avoid that some part of the interval $[0,1]$ be unreached by the distribution and that $\Pr(v=v_0) > 0$ for some v_0; i.e., we avoid in particular the case where v can only take a finite number of values, and in which the incomplete learning result established in Section III doesn't hold any more.)

Let F be the cumulative distribution function corresponding to density f. If the monopolist, starting from the information: $v \in I_0 = [0,1]$ and the c.d.f. F, ends up after t periods with the information: $v \in I_t = [x,y]$, then his posterior on $[x,y]$ at time t will be given by the c.d.f.:

$$F_{xy}(P) = \frac{F(P)-F(x)}{F(y)-F(x)} \quad , \text{ for } p \in [x,y] \quad .$$

Let $V_\delta(x,y)$ be the maximum intertemporal expected profit of the monopolist when he starts from the information $v \in [x,y]$ with c.d.f. F_{xy}.

<u>Lemma 1</u>: V_δ is the unique bounded solution of the Bellman equation:

$$\text{(B):} \quad V(x,y) = \max_{p\in[x,y]} \left[\frac{F(y)-F(p)}{F(y)-F(x)} (p+\delta\cdot V(p,y)) + \frac{F(p)-F(x)}{F(y)-F(x)} \delta \cdot V(x,p) \right] \quad .$$

The proof is identical to the one done in the uniform distribution case.

Corollary: V_δ is continuous and increasing in x and y .

Proof of Corollary: Here again the proof is quite similar to the one done in the uniform distribution case: We know by lemma 1 that V_δ is the unique fixed point on the Banach space B_0 (bounded functions on $[0,1]^2$) of the contraction mapping Π defined by (B). Therefore, it suffices to show that whenever V is continuous and increasing in x and y , $\Pi(V)$ inherits the same properties.

Let $X = F(x)$, $Y = F(y)$, $P = F(p)$, and $Q \in [0,1]$ such that: $P = (Y-X)Q + X$. We have:

$$\Pi(V)_{(x,y)} = \max_{Q\in[0,1]} [(1-Q)[F^{-1}((Y-X)Q+X) + \delta V(F^{-1}((Y-X)Q+X),y) + Q\delta V(X,F^{-1}((Y-X)Q+X)] = a(V,F,x,y,Q) .$$

For each $Q \in [0,1]$, $(x,y) \mapsto (Y-X)Q+X$ is increasing in x and y , and continuous. Given that F is continuous and increasing, and that V also is continuous and increasing in x and y by assumption, the mapping $(x,y) \mapsto a(V,F,x,y,Q)$ inherits the same property for all Q , and so does $\Pi(V) = \max_{Q\in[0,1]} a(V,F,.,.,Q)$.

This completes the proof. □

Lemma 2: Let $x_\delta = \frac{\bar{f}}{f+\underline{f}(1-\delta)}$; then $V(x,y) = \frac{x}{1-\delta}$ is the solution of the Bellman equation (B) on the set:

$$D_{x_\delta} = \{(x,y) \mid y/x \geq x_\delta\} .$$

(Therefore, from Lemma 1, $V_\delta(x,y) = \frac{x}{1-\delta}$ for $\frac{y}{x} \geq x_\delta$.)

Remark: In the uniform case, we have $\underline{f} = \bar{f} = 1$, so that: $x_\delta = \frac{1}{2-\delta}$ (Section II).

Proof of Lemma 2: Let $V(x,y) = \frac{x}{1-\delta}$ and let us calculate:

$$\max_{p\in[x,y]} \left[\frac{F(y)-F(p)}{F(y)-F(x)} \cdot \frac{p}{1-\delta} + \frac{F(p)-F(x)}{F(y)-F(x)} \cdot \delta \frac{x}{1-\delta} \right.$$

$$= \max_{p\in[x,y]} G(p) \quad .$$

We have: $G'(p) = \frac{1}{(1-\delta)(F(y)-F(x))} (F(y)-F(p)-f(p)\cdot(p-\delta x))$.

But: $F(y)-F(p)-f(p)(p-\delta x) \leq \bar{f}\ (y-p) - \underline{f}\ (p-\delta x)$

$$\leq \bar{f}\ y - [\bar{f} + \underline{f}\ (1-\delta)]x\ , \ \forall p \in [x,y] \quad .$$

So, for $x > \frac{\bar{f}}{\bar{f}+\underline{f}(1-\delta)}\, y = x_\delta \cdot y$, $G'(p) < 0$ for all $p \in [x,y]$ so that the maximum is obtained for $p = x$. But $G(x) = \frac{x}{1-\delta}$. This establishes the lemma. □

Lemma 4: $\exists\, k_\delta < 1$ such that $x \leq x_\delta \cdot y \Rightarrow p_1(x,y) \leq k_\delta \cdot y$, where $p_1(x,y)$ is the optimal price played by the monopolist in period 1 when he starts from the information: $v \in [x,y]$.

Proof of Lemma 4: Suppose $x \leq x_\delta \cdot y$ and $p_1 = p_1(x,y) \geq k_\delta \cdot y$, for all $k_\delta < 1$. Let's take (w.l.o.g.) $k_\delta \geq x_\delta$. (This is consistent with the inequality $k_\delta < 1$ since $x_\delta < 1$.)

Since: $p_1 \geq k_\delta \cdot y \geq x_\delta \cdot y$, we have: $V_\delta(p_1,y) = \frac{p_1}{1-\delta}$.

Then: $V_\delta(x,y) = \frac{F(y)-F(p_1)}{F(y)-F(x)} \cdot \frac{p_1}{1-\delta} + \frac{F(p_1)-F(x)}{F(y)-F(x)} \cdot \delta \cdot V_\delta(x,p_1)$.

Now, since $V_\delta(x,y) \geq V_\delta(x,p_1)$, we must have:

$$V_\delta(x,y)\cdot[F(y)-(1-\delta)F(x)-\delta F(p_1)] \leq (F(y)-F(p_1)) \cdot \frac{1}{1-\delta} \quad .$$

The one-shot monopoly profit starting from the information: $v \varepsilon [0,y]$ is:

$$\max_{p\varepsilon[0,y]} \frac{F(y)-F(p)}{F(y)} p \geq \max_{p\varepsilon[0,y]} \frac{\underline{f}}{\overline{f}\cdot y} (y-p)p$$

$$\geq \frac{\underline{f}}{4\overline{f}} \cdot y \ .$$

$$\Rightarrow V_\delta(x,y) \geq V_\delta(0,y) \geq \frac{y}{1-\delta} \cdot \frac{\underline{f}}{4\overline{f}} \ .$$

So we must have:

$$[(1-\delta)(F(y)-F(x))+\delta(F(y)-F(p_1))] \frac{\underline{f}}{4\overline{f}} y \leq (F(y)-F(p_1))p_1 \ .$$

Using $x \leq x_\delta \cdot y$ and letting $d = \underline{f}/\overline{f} \leq 1$, we get as a necessary condition:

$$\frac{d^2}{4} \cdot [(1-\delta)(1-x_\delta) + \delta(1-\frac{p_1}{y})] \leq (1-\frac{p_1}{y}) \frac{p_1}{y} \ .$$

The RHS is 0 when $\frac{p_1}{y} = 1$; the LHS is strictly positive when $\frac{p_1}{y} = 1$.

Therefore $\exists\ k_\delta < 1$ such that $\frac{p_1}{y} \leq k_\delta$. Q.E.D. □

Once we have these last two lemmas, the theorem stated in Section III and proved so far in the uniform case can immediately be generalized to the distributions $f(v)$ where $0 < \underline{f} \leq f(v) \leq \overline{f}$. It suffices to repeat the proofs of steps a and b, Section III.

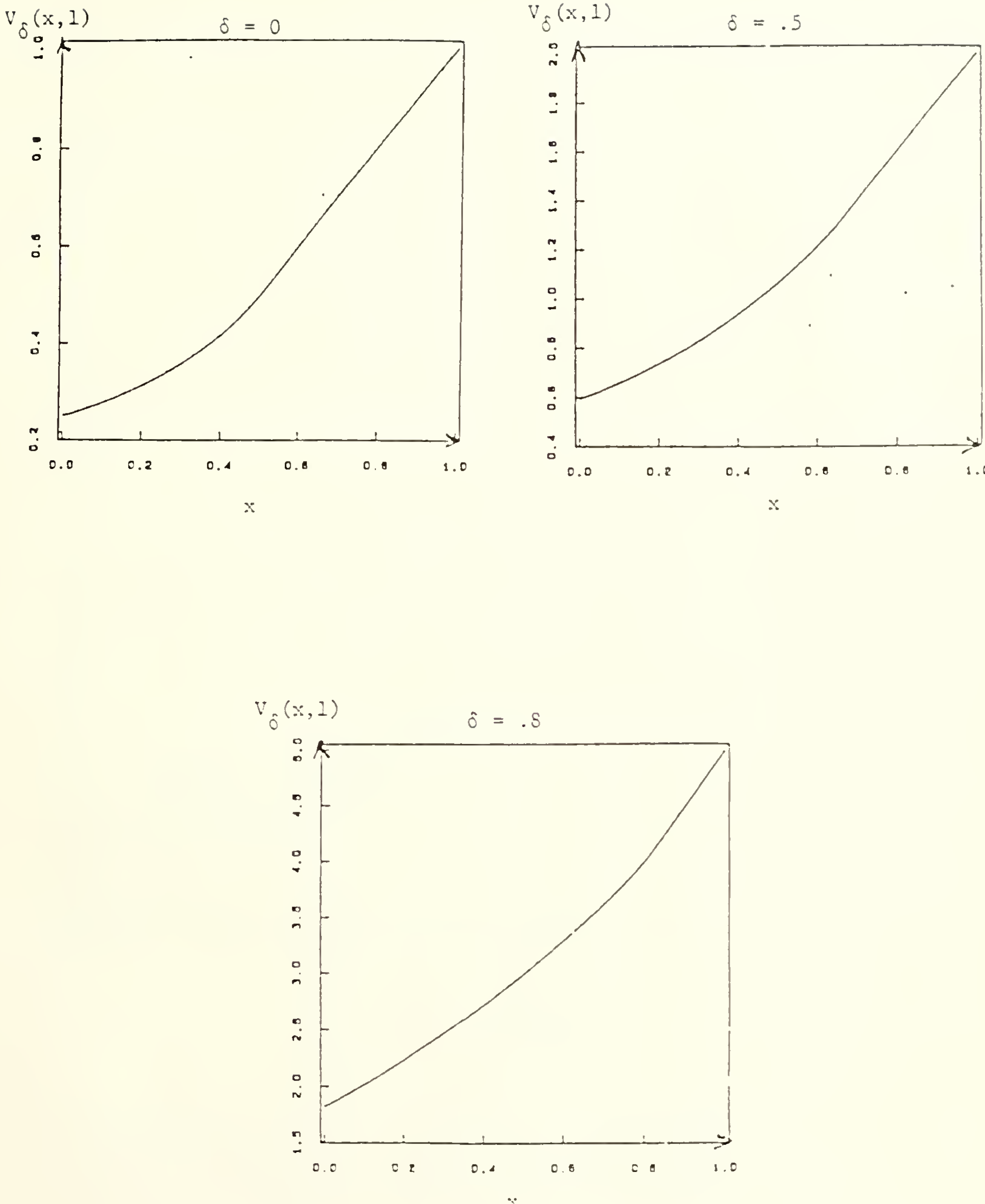


Figure 1

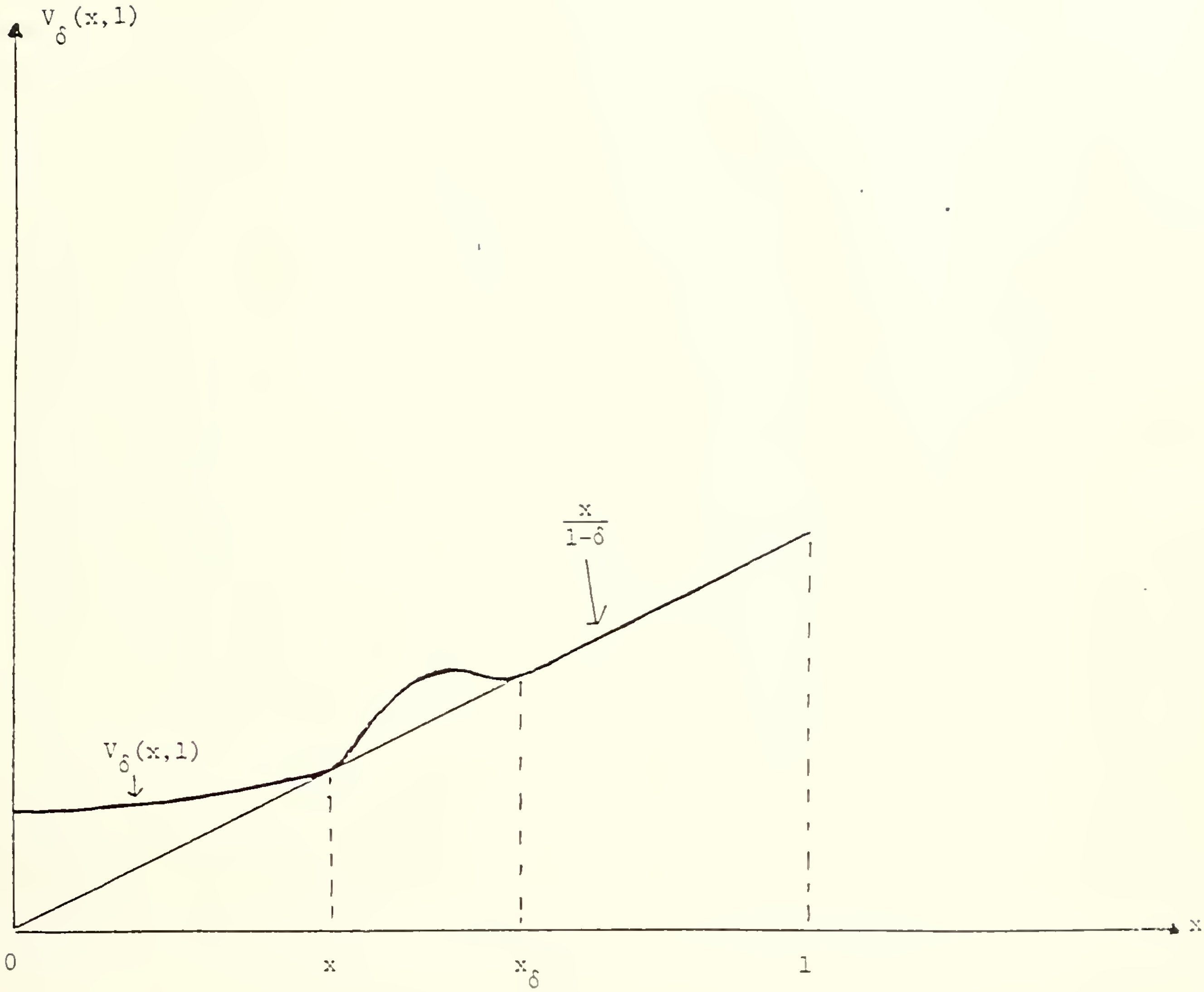


Figure 2

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

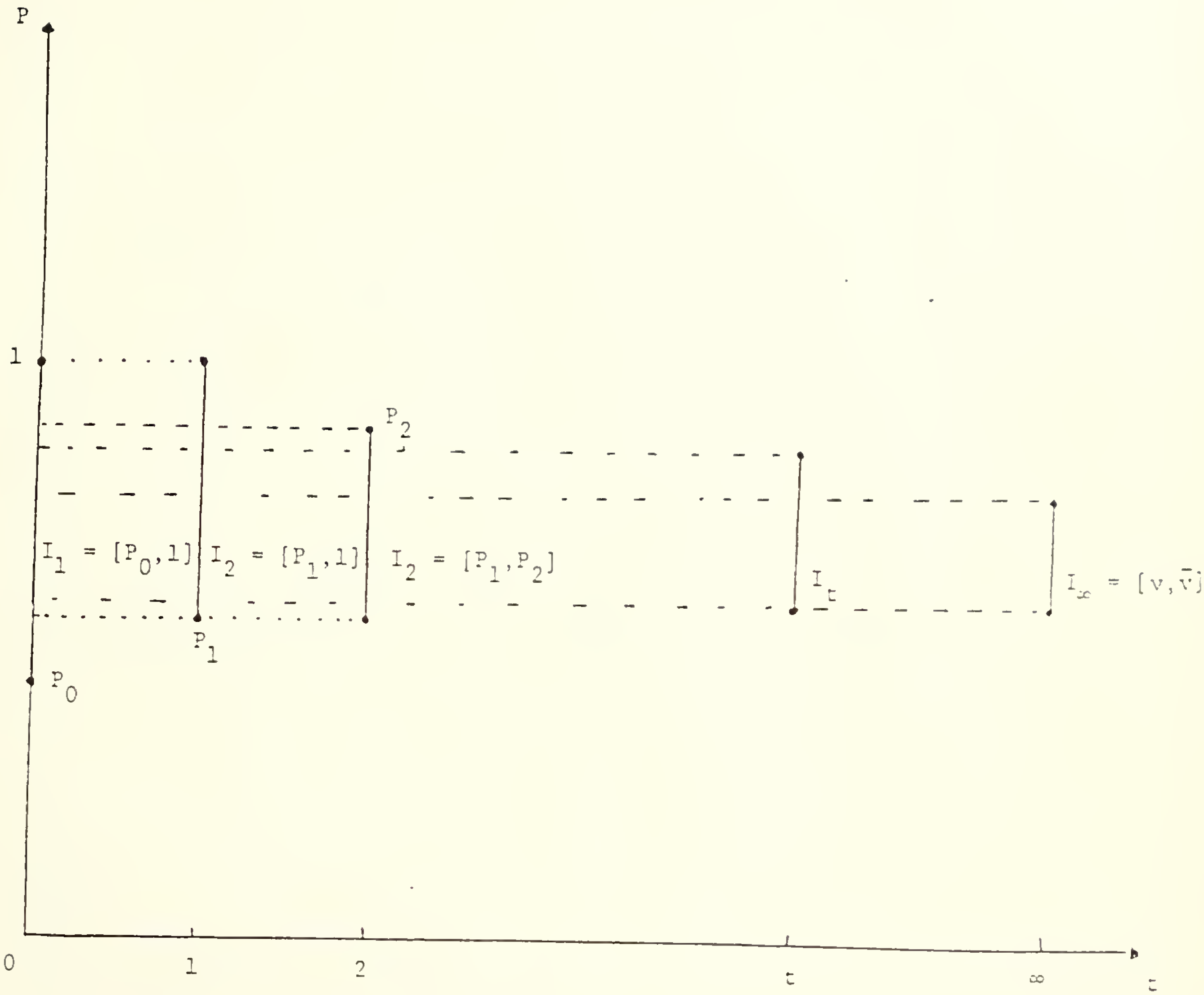
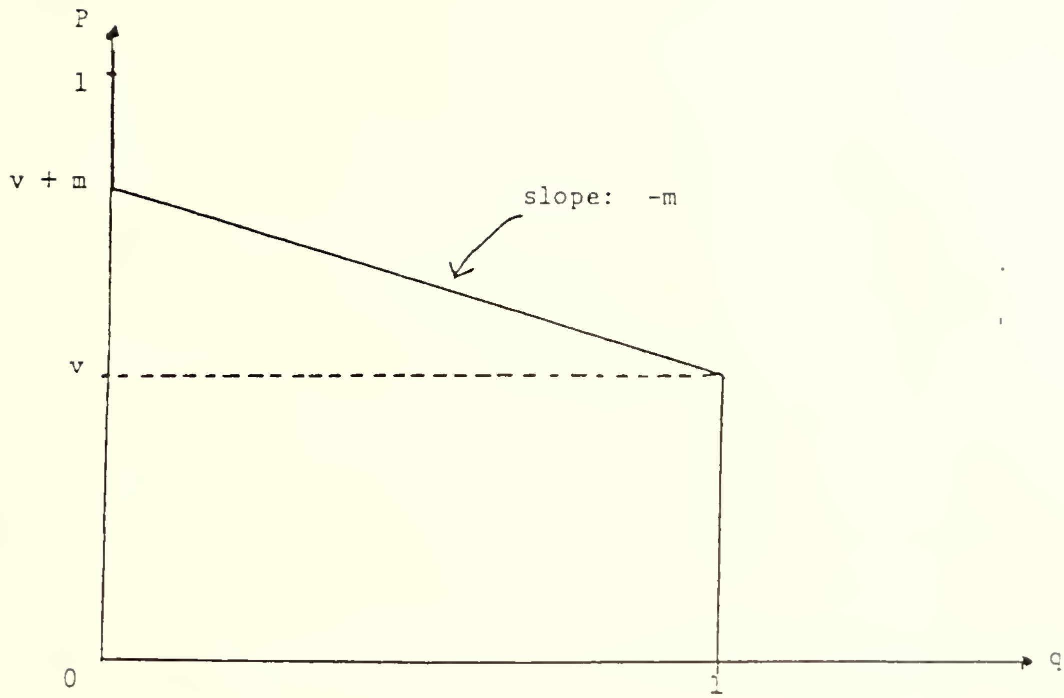


Figure 3



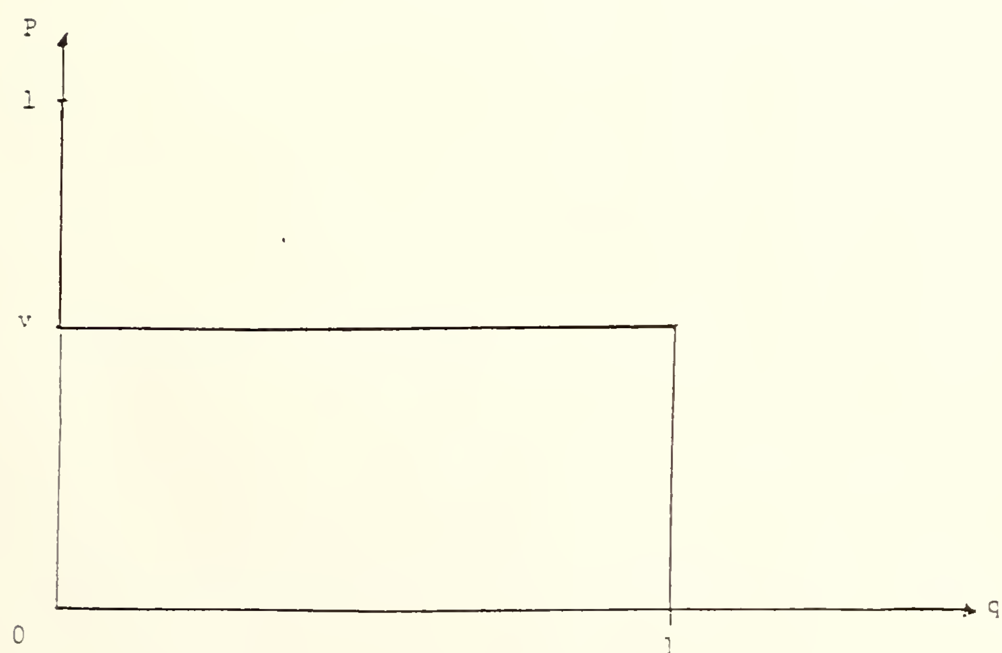


Figure 4

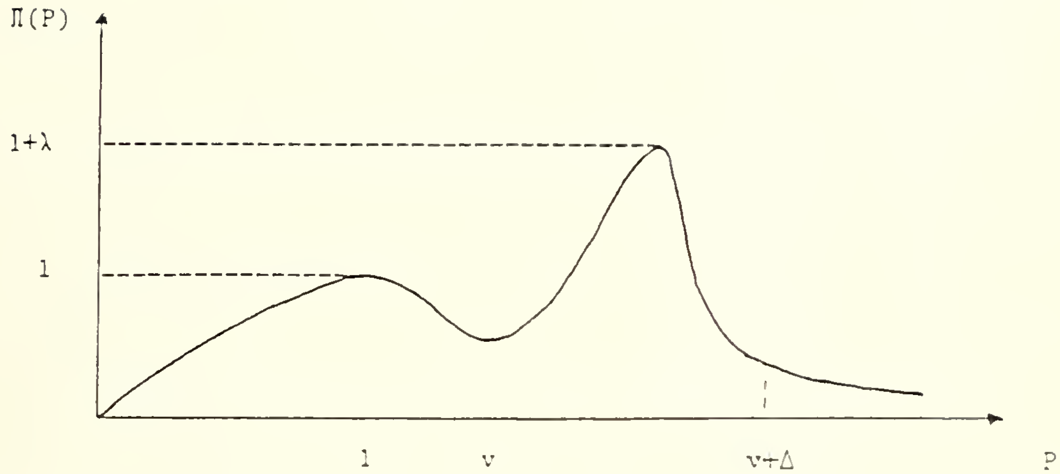


Figure 5

3-22-89