Discussion Paper No. 1015

LEARNING ABOUT VARIABLE DEMAND IN THE LONG RUN

by

Aldo Rustichini*

and

Asher Wolinsky**

September 1992

*Department of Economics, New York University, New York, New York 10003

**Department of Economics, Northwestern University, Evanston, Illinois 60208

LEARNING ABOUT VARIABLE DEMAND IN THE LONG RUN

Aldo Rustichini* and Asher Wolinsky**

Abstract

This paper studies the problem of a monopoly who is uncertain about the demand it faces and learns about it over time through its pricing experience. The demand curve facing the monopoly is not constant--it changes over time in a Markovian fashion. We characterize the monopoly's optimal policy and inquire how it differs from an informed monopoly's policy. It turns out that, even when the rate at which the demand varies is negligible, the stationary probability with which the monopoly's policy deviates from its informed counterpart is non-negligible, as long as the discount factor is below 1.

1. Introduction

This paper studies the problem of a monopoly who is uncertain about the demand it faces and learns about it over time through its pricing experience. The existing literature on this problem includes the work of Aghion, Bolton, Harris and Jullien (1991), Mclennan (1984,1986), Easely and Kiefer (1989) and Kiefer(1987). The common basic problem, which the first three contributions share, is one in which the demand curve is fixed and through learning the monopoly narrows down its initial uncertainty about it. The new angle that Kiefer (1987) and the present paper add to this literature is that the demand curve facing the monopoly is not constant--it changes over time in a Markovian fashion. Thus, even after the monopoly has learnt the state of demand, the demand is liable to change and the monopoly has to check occasionally whether it has changed. In Kiefer(1987) the set-up is significantly more ambitious than ours and consequently it does not obtain many analytic results. We focus, in turn, on a particularly simple example which allows us to accommodate the added complexity and to obtain relatively sharp insights into the optimal monopoly's pricing policy.

In our example, the state of demand follows a symmetric Markov chain and in the different states the demand curves are inelastic unit demands which differ in the maximum price buyers are willing to pay. After characterizing the monopoly's optimal policy, we examine the extent of learning--how different the monopoly's policy is from what it would be if the monopoly knew the state of demand. It turns out that, even when the rate at which the demand varies is negligible, the stationary probability with which the monopoly's policy deviates from its informed counterpart is non-negligible as long as the discount factor is below 1.

2. Model and Analysis

The demand for the monopoly's product varies stochastically over time. Time is divided into discrete periods labelled $t=1,2,\ldots$. In each period the monopoly faces a unit demand with reservation price d^t. At the beginning of the period, before it knows d^t, the monopoly quotes a price p^t. Thus, if $p^t \le d^t$ the monopoly will sell a unit, and if $p^t > d^t$ it will sell nothing. Let I^t record whether or not there has been a sale at period t. That is, $I^t=1$ if $p^t \le d^t$ and 0 otherwise. The maximum demand price, d^t, can assume two values: 1 and $D>1$. It follows a Markov process with transition probabilities

$$\mathrm{Prob}[d^{t+1}=1 \mid d^t=D]=\mathrm{Prob}[d^{t+1}=D \mid d^t=1]=\alpha .$$

Let w^t denote the probability with which the monopoly believes that $d^t=1$. Thus, w^1 is the prior and subsequently it evolves as follows

$$w^{t+1} = \begin{cases} \alpha & \text{if } 1<p^t \le D \text{ and } I^t=1 \\ 1-\alpha & \text{if } 1<p^t \le D \text{ and } I^t=0 \\ (1-\alpha)w^t+\alpha(1-w^t) & \text{if } p^t=1 \end{cases}$$

Given a price sequence $\{p^t\}$ and a sequence of demand realizations $\{d^t\}$, the monopoly's discounted profit is $\sum \delta^t p^t I^t$, where $\delta<1$ is the discount factor. At the beginning of period t the monopoly knows the history $h^t=(p^1,I^1),\ldots,(p^{t-1},I^{t-1})$. Its problem is to choose a pricing policy $p^t(h^t)$ so as to maximize $E[\sum \delta^t p^t I^t]$.

Note that w^t is determined by the prior, w^1, and the history, h^t. The following claim restates the well known fact that this monopoly problem has a stationary optimal solution which depends only on w^t (see, e.g., Derman (1970)), and that the dependence on w^t takes a particularly simple form.

Claim 1: The optimal policy is characterized by a cutoff belief W. If $w^t \leq W$, $p^t = D$. If $w^t > W$, $p^t = 1$.

Let N be the smallest nonnegative integer such that, if $w^t = 1-\alpha$ and $p^t = \ldots = p^{t+N-1} = 1$, then $w^{N+t} \leq W$; let $N = \infty$ if there is no such integer. That is, N is the number of times the seller quotes p=1 after a price offer $p_{t-1} = D$ was rejected. Note that N essentially characterizes the optimal policy, except perhaps in the first few steps. That is, after a p=D was rejected the monopoly will quote p=1 for N times in a row, then quote p=D until rejected, and so on.

The following analysis, which is summarized by Claim 2 below, characterizes the optimal policy (i.e., the optimal N) in terms of the parameters. Let $\psi_N = \text{Prob}[d^{t+N+1} = 1 | d^t = 1]$. Standard calculations (see, e.g., Feller(1968)) yield

(1) $$\psi_N = [1 + (1-2\alpha)^{N+1}]/2.$$

Let Z denote the expected future profit, evaluated at the end of a period in which a price D was rejected. Thus, Z is the expected profit generated by a sequence of prices which starts with N consecutive periods of p=1.

Let Y denote the expected future profit, evaluated at the end of a period in which a price D was accepted. Thus, Y is the expected profit generated by a sequence which starts with p=D in the next period.

(2) $$Z = (\delta - \delta^{N+1})/(1-\delta) + \delta^{N+1}\psi_N Z + \delta^{N+1}(1-\psi_N)(D+Y)$$

(3) $$Y = \delta[(1-\alpha)(D+Y) + \alpha Z]$$

The solution to system (2)-(3) is

$$(4) \qquad (1-\delta)Z = \frac{2[1-\delta(1-\alpha)](\delta-\delta^{N+1})/(1-\delta) + \delta^{N+1}[1-(1-2\alpha)^{N+1}]D}{2[1-\delta(1-\alpha)](1-\delta^{N+1})/(1-\delta) + \delta^{N+1}[1-(1-2\alpha)^{N+1}]}$$

The monopoly's optimal N is of course the one that maximizes $(1-\delta)Z$. Let $\mathrm{NUM}\{d[(1-\delta)Z]/dN\}$ denote the numerator of the derivative $d[(1-\delta)Z]/dN$.

$$(5) \qquad \mathrm{NUM}\{d(1-\delta)Z/dN\} = \delta^{N+1}[1-\delta(1-\alpha)]\{2[1-\delta(1-\alpha)]\log\delta + (D-1)\delta^{N+1}(1-2\alpha)^{N+1}\log(1-2\alpha) + (D-\delta)[\log\delta-(1-2\alpha)^{N+1}\log((1-2\alpha)\delta)]\}/(1-\delta)$$

<u>Claim 2</u>:

If $D\le 1+[1-\delta(1-2\alpha)]$, then $N=\infty$.

If $1+[1-\delta(1-2\alpha)]<D<2[1-\delta(1-2\alpha)]\log\delta/[2\alpha\log\delta-(1-\delta)(1-2\alpha)\log(1-2\alpha)]$, then N will be one of the integers which are closest to the unique solution to the equation $\mathrm{NUM}\{d(1-\delta)Z/dN\}=0$.

If $D\ge 2[1-\delta(1-2\alpha)]\log\delta/[2\alpha\log\delta-(1-\delta)(1-2\alpha)\log(1-2\alpha)]$, then $N=0$.

<u>Proof</u>: Let $T(N)$ denote the term in the curly brackets on the RHS of (5). Note that $T'(N)<0$. Therefore, there are three possible cases: (i) if $T(N)>0$, for all N, then $Z(1-\delta)$ is monotonically increasing in N, so that $N=\infty$ is the optimum; (ii) if $T(0)\le 0$, then $Z(1-\delta)$ is monotonically decreasing in N, so that $N=0$ is the optimum; (iii) if $T(N)=0$, for some $0<N<\infty$, then this N is the maximizer of $Z(1-\delta)$, and since $T'(N)<0$ the integer maximizer is one of the two integers closest to it.

Now, $T(N)>0$, for all N, iff $T(\infty)=-2[1-\delta(1-\alpha)]\log\delta+(D-\delta)\log\delta\ge 0$, iff $D\le 1+[1-\delta(1-2\alpha)]$.

$T(0) = -2[1-\delta(1-2\alpha)]\log\delta + D[2\alpha\log\delta-(1-\delta)(1-2\alpha)\log(1-2\alpha)] \le 0$ iff $D\ge 2[1-\delta(1-2\alpha)]\log\delta/[2\alpha\log\delta-(1-\delta)(1-2\alpha)\log(1-2\alpha)]$.

For D in the middle range $T(0)>0$ and $T(\infty)<0$, so there exists a unique N for which $T(N)=0$. QED

Thus, given δ and α, there are three types of optimal policy, depending on the relative size of D. When D is sufficiently small or sufficiently large the monopoly will quote always the same price, 1 or D respectively. When D is in the intermediate range, the optimal policy involves price changes whose frequency depends on N.

3. The long run behavior

In implementing the optimal policy the monopoly will make two kinds of error. In some periods it will charge $p^t=1$ when $d^t=D$, while in others it will charge $p^t=D$ when $d^t=1$. The stationary probabilities of these two types of error capture the frequency with which these errors are made in the long run, and hence provide some measure of the extent of learning associated with the optimal policy. In what follows we examine these probabilities, particularly when the demand changes infrequently (α small). As noted in Claim 2, there are three types of optimal policy, depending on the relative size of D. When D is sufficiently small or sufficiently large, the monopoly will quote only one price. Hence, there will be only one kind of error and its stationary probability will be the stationary probability of $d^t=D$ or $d^t=1$ as might be the case. Therefore, the only interesting cases for analysis are when D is in the intermediate range, where the optimal policy involves price changes, and we shall focus below on this case.

There are four possible combinations of (p^t,d^t): (1,1),(1,D),(D,1) and (D,D). Let Π denote the stationary probabilities of these combinations. I.e.,

$$\Pi(p,d)=\lim_{T\to\infty}\#\{t:\ (p^t,d^t)=(p,d),\ t\leq T\}/T.$$

The following claim expresses the long run probabilities of the monopoly's errors, $\Pi(1,D)$ and $\Pi(D,1)$, in terms of the parameters and N.

<u>Claim 3</u>:
$$\Pi(1,D) = \frac{(1-2\alpha)[1-(1-2\alpha)^N]-2\alpha N}{2[2(N+1)\alpha+1-(1-2\alpha)^{N+1}]}; \quad \Pi(D,1) = \frac{2\alpha}{2(N+1)\alpha+1-(1-2\alpha)^{N+1}}$$

<u>Proof</u>: Initially, let us merge the states $(p^t,d^t)=(1,1)$ and $(p^t,d^t)=(1,D)$ into one state denoted $(1,.)$, and look at the resulting three state Markov Chain. Its transition matrix, P, is

	(1,.)	(D,D)	(D,1)
(1,.)	0	$1-\psi_N$	ψ_N
(D,D)	0	$1-\alpha$	α
(D,1)	1	0	0

The stationary probabilities of this Markov chain, $\pi=[\pi(1,.),\pi(D,D),\pi(D,1)]$, are the solution to the system $\pi P=\pi$. Thus,

$\pi(1,.) = \pi(D,1) = \alpha/(2\alpha+1-\psi_N)$; $\pi(D,D) = (1-\psi_N)/(2\alpha+1-\psi_N)$.

Note, however, that in terms of the periods of the original model a visit to state $(1,.)$ lasts N periods, while visits to the other two states last one period each. After adjusting for this fact, the stationary probabilities are

(6) $\Pi(1,.)=N\alpha/[(N+1)\alpha+1-\psi_N]$; $\Pi(D,D)=(1-\psi_N)/[(N+1)\alpha+1-\psi_N]$;

$\Pi(D,1)=\alpha/[(N+1)\alpha+1-\psi_N]$.

Now, to split state $(1,.)$ into $(1,D)$ and $(1,1)$ let

$K(1,n)=E[\#\{t: d^t=D,\ 1<t\leq n+1\}|d^1=1]$ and $K(D,n)=E[\#\{t: d^t=D,\ 1<t\leq n+1\}|d^1=D]$.

Note that

(7) $\Pi(1,D)=\Pi(1,.)K(1,N)/N$; $\Pi(1,1)=\Pi(1,.)[N-K(1,N)]/N$

To get an explicit formula for $K(1,N)$, note that $K(1,1)=\alpha$, $K(1,n+1)=\alpha+\alpha K(D,n)+(1-\alpha)K(1,n)$ and $K(D,n)+K(1,n)=n$. It follows that $K(1,n)=\{(1-2\alpha)[1-(1-2\alpha)^n]-2\alpha n\}/4\alpha$. Upon substituting the expression for ψ_N from (1) into (6) and then substituting the result and $K(1,N)$ into (7), we get the desired formulae. QED

If the changes in demand are frequent, it is obvious that, for a non-negligible fraction of the time, the monopoly's prices will not be in perfect accordance with the state of demand. It is less obvious what happens when the demand changes infrequently. That is, how do $\Pi(1,D)$ and $\Pi(D,1)$ behave when α is arbitrarily close to 0.

Claim 4:

$$\lim_{\alpha\to0}\Pi(1,D) = \frac{2(1-\delta)+(D-\delta)\log[1-2(1-\delta)/(D-\delta)]}{4(1-\delta)-2(D-\delta)\log[1-(1-\delta)/(D-\delta)]}; \quad \lim_{\alpha\to0}\Pi(D,1) = 0.$$

Proof: As α approaches 0, it follows from the RHS of (5) that

$$\lim_{\alpha\to0}\{-2(1-\delta)+(D-\delta)[1-(1-2\alpha)^{N+1}]=0.$$

Hence, $1-2(1-\delta)/(D-\delta)=\lim_{\alpha\to0}(1-2\alpha)^N=e^{-2\lim(N\alpha)}$.

Therefore, $\lim_{\alpha\to0}(N\alpha)=-\log[1-2(1-\delta0/(D-\delta)]/2$.

Taking the limits of the expressions given in the statement of Claim 3, using the above expressions for $\lim_{\alpha\to0}(1-2\alpha)^N$ and $\lim_{\alpha\to0}(N\alpha)$, we get the $\lim\Pi(1,D)$ and $\lim\Pi(D,1)$ formulae in the statement of the claim. QED

4. Discussion

Learning in the long run

Notice that, for $\delta<1$, $\lim_{\alpha\to0}\Pi(1,D)>0$. This means that, even when the frequency of changes in demand is negligible, there is a non-negligible fraction of the time during which the monopoly's price deviates from what it would be if the monopoly knew the true state. This is the result of balancing two opposing forces. For a given N, the smaller is α the smaller will be the probability of error $\Pi(1,D)$. But a small α induces the monopoly to "check" less frequently the state of demand (i.e., N is large), so that once the state of demand switches from 1 to D, it might take longer for the monopoly to find this out. The latter force also explains why the probability of the other type of error, $\Pi(D,1)$, vanishes. That error occurs when the monopoly raises its price to D to test whether the demand has changed and when the state of demand changes while the monopoly is charging D. A large N diminishes occurrences of the former type while a small α diminishes occurrences of the latter type.

These observations are better understood by comparing them with what happens when the demand does not vary at all, i.e., $\alpha=0$. When the demand is fixed, the monopoly's optimal pricing policy depends on its initial probability belief, $w^1=\text{Prob}\{d=1\}$, where d is the constant state of demand. If the monopoly decides to learn about d, it chooses $p^1=D$. This is optimal if the expected benefit of learning about demand, at the risk of possibly missing the first period's sales, exceeds the profit of charging the riskless price p=1. That is, if

$$(1-w^1)D/(1-\delta) + w^1\delta/(1-\delta) \geq 1/(1-\delta).$$

Otherwise, the monopoly chooses p=1 in perpetuity and never learns the true d. In other words, the monopoly learns the true demand if $w^1 \leq (D-1)/(D-\delta)$. In particular, if $w^1=1/2$ the monopoly learns the true demand if $1/2 \leq (D-1)/(D-\delta)$. Now, in the variable demand case, due to the symmetry, the stationary probability of $d^t=1$ is 1/2. That is, if the monopoly quotes p=1 for sufficiently long time, the belief $w^t=\text{Prob}\{d^t=1\}$ will approach 1/2. If α is sufficiently small, the benefits of learning about demand are close to what they are in the fixed demand case, so that if $1/2<(D-1)/(1-\delta)$, the monopoly will switch after a while to D, and if $1/2>(D-1)/(1-\delta)$ the monopoly will charge only p=1 in the long run. Regarding learning, the important difference between the cases of α arbitrarily close to 0 and $\alpha=0$ (with $w^1=1/2$) is that when $1/2<(D-1)/(1-\delta)$ the learning is perfect in the latter case, but is imperfect in the former case in the sense that the long run probability of a pricing error, $\Pi(1,D)$, is non-negligible.

The cost of learning is the possible loss of a unit profit, while the benefit depends on D and δ and is clearly increasing in both. Observe from the formula given in Claim 4 that $\lim_{\delta\to 1}\lim_{\alpha\to 0}\Pi(1,D)=0$. That is, when the demand changes very infrequently and the cost of learning is negligible next to the expected benefit, the probability of a pricing error is negligible.

Further remarks on the assumptions

The relatively sharp characterization of the long run behavior of this model owes of course to some of its special assumptions. The important ones are the finiteness of the underlying states of demand and the finiteness of the relevant actions (prices). The specific assumptions that there are only two states, only two relevant prices and that the transition between the states is symmetric, simplify the exposition but are not crucial. However, the finiteness of the underlying states of demand and the set of relevant prices do seem important in a way that does not allow to speculate on what the results might be in the case with a continuum of states and actions. The role of the finiteness of the set of relevant prices, which is implied by the simple nature of the demand, is that it introduces some rigidity into the seller's learning opportunities. Specifically, it does not allow the seller to learn from small variations in price, so learning always involves some non-negligible cost.

We should emphasize that we do not feel too apologetic about these assumptions. They mean that this model addresses a situation in which the decision maker faces only a few discernible options and the choice of one rather than the other involves some non-negligible difference in payoff. The other model, with continuum of states and actions, is also interesting and will be useful to analyze (though it is probably substantially more complicated), but it addresses a fundamentally different situation and does not necessarily provide a more appropriate model for the pricing behavior we discuss.

References

Aghion, P., Bolton, P., Harris, C. and B. Jullien (1991), "Optimal Learning by Experimentation," Review of Economic Studies, 58, 621-654.

Derman, C. (1970), Finite State Markovian decision Processes, New York: Academic Press

Easely, D. and N. Kiefer (1988), "Controlling a Stochastic Process with Unknown Parameters," Econometrica, 56, 1045-1064.

Feller, W. (1968), An Introduction to Probability and its Application, (3rd edition), Wiley, New York.

Kiefer, N. (1987), "A Model with Optimal Learning and Obsolescence of Information," mimeo, Cornell University.

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