



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

## Adaptative Learning Models of Consumer Behaviour (first version)

### Citation for published version:

Hopkins, E 2003 'Adaptative Learning Models of Consumer Behaviour (first version)' ESE Discussion Papers, Edinburgh School of Economics, University of Edinburgh.  
<<http://ideas.repec.org/p/edn/esedps/80.html>>

### Link:

[Link to publication record in Edinburgh Research Explorer](#)

### Document Version:

Peer reviewed version

### Publisher Rights Statement:

© Hopkins, E. (2003). Adaptative Learning Models of Consumer Behaviour (first version). (ESE Discussion Papers). Edinburgh School of Economics, University of Edinburgh.

### General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

### Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



# Adaptive Learning Models of Consumer Behaviour

Ed Hopkins\*  
Department of Economics  
University of Edinburgh  
Edinburgh EH8 9JY, UK

February, 2002  
Preliminary

## Abstract

This paper applies recent advances in the theory of learning to the analysis of consumer behaviour. The working assumption is that while sellers are rational in the traditional sense, consumers are boundedly rational. The differences in outcomes for search goods and experience goods are investigated. In the latter case, if consumers fail to take into account that information is only partial, they can become locked into the habit of purchasing inferior goods. Surprisingly, however, prices are lower than when information is complete. Firms have an incentive to offer lower prices to prevent consumers becoming locked into their rival's product.

---

\*E.Hopkins@ed.ac.uk, <http://www.ed.ac.uk/~ehk>

# 1 Introduction

Adaptive learning theory attempts to describe the behaviour of agents faced with repeated decision problems by assuming they use simple learning rules. The principal question has been under what circumstances do agents learn to choose an optimal action and when do they become locked into an inferior choice.<sup>1</sup> Up to now, this has largely been explored in the context of abstract environments. However, the same issues are present in the empirical analysis of consumer behaviour, where marketing analysts attempt to explain the choices of consumers over different brands. The terminology employed is different. Agents choose over “brands” instead of “actions”, or they may have a desire for “variety”, rather than a tendency to “experiment”. The underlying models are often similar.

In particular, the familiar division of products into *search* and *experience* goods has an exact parallel in the field of adaptive learning. Rustichini (1999) describes the situation where an agent obtains information on the payoff of only the action chosen as “partial” information. This corresponds to a consumer buying experience goods, as the consumer only discovers the qualities of the good she actually chooses, and not about the other brands she leaves in the store. When the agent sees each period the payoff to all actions, Rustichini terms this “full” information. This is similar to the situation with search goods, as a consumer can estimate the qualities of all brands, not just the one he purchases. What Rustichini finds is that under both partial and full information, depending on the learning rule adopted, there is the possibility of an agent becoming locked into an inferior choice.

In this paper, I extend the model of Chintagunta and Rao (1996), which examines a dynamic duopoly where consumers learn over time, to consider a wider variety of learning rules and information treatments. I find that under full information, learning is complete and consumers eventually have correct beliefs about the quality of the goods on offer. However, when goods are experience goods, if consumers fail to take into account that information is only partial, they can become locked into inferior choices. Surprisingly, this leads to lower prices as firms compete to lock consumers into their products. If, in contrast, consumers adopt a learning rule that is optimal under partial information, asymptotically prices will be the same as under complete information. The difference is that under partial information, even with an optimal learning rule, firms can influence consumer opinion in the short run.

What is distinctive about this approach is the mixture of rational behaviour by sellers and boundedly rational behaviour by buyers. The recent literature on adaptive learning has largely focussed on abstract exogenous environments. Exceptions that deal with market models have concentrated on sellers (Capra et al, 2000; Hopkins and Seymour, 2000). However, there are reasons to believe that models of bounded rationality may have a better fit to consumer behaviour. As Erdem et al (1999) point out, consumers may be prone to various

---

<sup>1</sup>References in this field include Rustichini (1999), Börgers and Sarin (2000), Heller and Sarin (2001), Sarin and Vahid (1999).

cognitive biases. One well known problem is spurious differentiation, when consumers form an incorrect belief that one brand or product is superior. For example, one report of the Consumer Union (1991) found that less than 50% of consumers sample were able in blind tasting to identify the soft drink brand they claimed to prefer. Yet much of the recent work on consumer learning about product quality takes the form of consumers making highly sophisticated inferences from the strategic behaviour of firms.<sup>2</sup>

The situation to be modelled can be thought of as a consumer going on a regular basis to a supermarket to buy a grocery item and choosing between two competing brands. This type of decision has several aspects which I would like to emphasise. First, the prices for the competing brands are usually clearly marked on the shelves. Thus, the learning the consumer has to undertake is not about prices or their distribution. However, the goods in question are typically experience goods. One has to take them home and consume them before their quality is known. Second, quality in this context is very often subjective, for example, whether a food product tastes good. Third, because each successive purchase decision is relatively unimportant to an individual consumer, a model of boundedly-rational behaviour may explain actual choices well.

This approach has already been taken in marketing. As noted above, marketing analysts are interested in how consumers choose and how their choices change over time. For example, a stochastic model of habit-persistence and variety seeking is explored in Seetharaman and Chintagunta (1998). Chintagunta and Rao (1996) go further in that they analyse the optimal seller response to a representative consumer who learns adaptively. What is a distinctive feature relative to the economics literature on learning is that this work is also empirical. These models are fitted to data on actual prices, sales and consumer purchases.

In the economics literature, there has been some interest in markets where not all consumers switch to the lowest price seller because there are switching costs associated with changing supplier. This situation has been examined in the context of dynamic duopoly by Padilla (1995), To (1996) and Chen and Rosenthal (1996). This compares with the current situation where consumers may be slow to switch between sellers, not because of direct costs, but because they respond adaptively to changes in prices.

What these models have in common is that, while sellers compete on prices, equilibrium prices are above competitive levels. This is because consumers switch between sellers only slowly. This emphasises that the speed and success of learning by consumers will determine the competitiveness of the market. More recently, Erev and Haruvy (2000) have considered the implications of adaptive learning for pricing policy. They find that the variability of prices may slow consumer learning and therefore reduce competitiveness. What the current paper attempts is similarly to link the success of learning to the amount and type of information available.

---

<sup>2</sup>See, for example, Bagwell and Ramey (1994).

## 2 Models of Adaptive Learning

Models of learning have been employed both to explain behaviour in games and in single person decision making. There is now considerable evidence that it can explain actual choice behaviour (Erev and Roth, 1998; Camerer and Ho, 1999; Erev and Barron, 2001). In games, payoffs are determined by the choices of one's opponents and in decision problems, by an exogenous random process, but in both cases, learning rules have three components. First, a decision maker is endowed with propensities (alternative terms are assessments or weights or scores), one for each of the possible actions in her action set. Second, there is a choice rule that chooses an action as a function of current propensities. A general principle is that actions with higher propensities are chosen with higher probability. Finally, there is an updating rule, which then changes the propensities in response to payoffs.

The propensities of a representative agent at time  $t$  we denote as  $\theta_t = (\theta_{1t}, \theta_{2t}, \dots, \theta_{nt}) \in \mathbb{R}^n$ , when the agent must choose from  $n$  actions. At each point in time, the agent chooses action  $i$  with probability  $x_{it}$ . The two choice rules that have attracted the most attention are the reinforcement learning rule

$$x_{it} = \frac{\theta_{it}}{\sum_{j=1}^n \theta_{jt}} \quad (1)$$

and the logit or exponential rule

$$x_{it} = \frac{\exp(\beta\theta_{it})}{\sum_{j=1}^n \exp(\beta\theta_{jt})}. \quad (2)$$

In the exponential choice rule, the parameter  $\beta$  represents the degree of optimisation. At high levels of  $\beta$  the agent will choose the action with the highest propensity with very high probability.

There are equally two prominent updating rules. The first takes the following form. If at time  $t$ , an agent takes action  $i$  and receives a payoff of  $u_{it}$ , then

$$\begin{aligned} \theta_{it+1} &= (1 - \delta)\theta_{it} + u_{it} \\ \theta_{jt+1} &= (1 - \delta)\theta_{jt}, \end{aligned} \quad (3)$$

where  $\delta$  is a “recency” parameter ( $\delta = 1$  implies no memory,  $\delta = 0$  implies perfect memory). In other words, this rule responds only to realised payoffs. Only the propensity corresponding to the action taken is updated. The payoff to other actions at that time was not observed. Rustichini (1999) refers to this as “partial” information. In contrast, “full” information is where the agent is able to see the payoffs to all actions at each time period. She is therefore able to update all propensities. For example, one can assume that

$$\theta_{it+1} = (1 - \delta)\theta_{it} + \delta u_{it} \text{ for } i = 1, 2, \dots, n. \quad (4)$$

The properties of these different learning rules in games are investigated in Hopkins (2000). Erev and Roth (1998) and Camerer and Ho (1999) both show how these models can be used to explain data from experimental games. In the context of single person decision problems, Rustichini (1999) finds the following. Suppose the payoffs to the actions are determined by a stationary probability distribution. Label the action with the highest expected return under this distribution, action 1. Then, by Propositions 3.2 and 5.2 (Rustichini, 1999), we have the following.

**Proposition 1** *Consider either the reinforcement rule (1) with partial information or the exponential choice rule (2) with full information and with  $\delta = 0$ . Then, almost surely  $\lim_{t \rightarrow \infty} x_{1t} = 1$ , that is, the probability placed on the action with the highest expected return goes to 1.*

But, importantly as Proposition 3.4 of Rustichini demonstrates, optimality is not always achieved.

**Proposition 2** *If the exponential choice rule (2) is used with the partial information updating rule with  $\delta = 0$ , then as  $t$  goes to infinity, any one action will be chosen with probability close to one.*

That is, the exponential choice rule in a situation of partial information can be interpreted as a form of overconfidence. With a high value of  $\beta$  the action that seems the best will be chosen with a high probability. Therefore, under partial information, the agent may never find out that another action would actually give a higher payoff.

### 3 A Model of Dynamic Duopoly

In this section, the dynamic duopoly model of Chintagunta and Rao (1996) (hereafter, “CR”) with learning by consumers is introduced. We will go on to consider a larger number of learning specifications than in the original paper. There are two competing firms, who each produce the same product at constant zero marginal cost. There is some ambiguity in this model as to whether these two goods are homogenous or differentiated. What is assumed is that consumers may possess a belief that one good is better than the other. For simplicity, we will initially consider a single representative consumer, an assumption that will be relaxed later. This consumer has goodwill for two sellers equal to  $(\theta_{1t}, \theta_{2t})$ . Each  $\theta_{it}$  can be thought of as the consumer’s estimate of quality of the  $i$ th brand. At each point in time the consumer seeks to buy one unit of the good.

The two firms compete on prices in continuous time. At any given time  $t$ , their prices can be written  $p_t = (p_{1t}, p_{2t})$ . The consumer uses a decision rule of the logit form. That is, the probability of purchasing from the  $i$ th seller is

$$x_{it}(\theta, p) = \frac{\exp(\beta\theta_{it} + \gamma p_{it})}{\sum_{j=1}^2 \exp(\beta\theta_{jt} + \gamma p_{jt})} \quad (5)$$

where  $\beta$  is a parameter measuring the sensitivity to goodwill, and  $\gamma$  measures sensitivity to prices.<sup>3</sup> It will be convenient to assume that  $\beta > 0$  and  $\gamma < 0$ . Duopoly with logit demand functions has been studied in the industrial organisation literature (see for example, Caplin and Nalebuff, 1991).

The consumer's goodwill for the two firms will change over time. If she consumes good  $i$  at time  $t$ , then she receives a utility of  $u_{it}$ . A *learning rule* in this context will be a way of updating goodwill  $\theta$  in response to this experience, giving rise to an equation of motion

$$\dot{\theta}_i = f_i(\theta, x(\theta, p), u). \quad (6)$$

In this paper, I consider three distinct functional forms, each representing different behavioural assumptions, and each having differing predictions.

Given the assumption of a representative consumer, a firm's instantaneous profits will be  $p_{it}x_{it}$ . Each firm attempts to maximise

$$\int_{t=0}^{\infty} e^{-rt} p_{it}x_{it}(\theta, p) dt \text{ subject to } \dot{\theta} = f(\theta, x(\theta, p), u). \quad (7)$$

This in turn gives rise to a current-value Hamiltonian for each firm

$$H_i = p_{it}x_{it}(\theta, p) + \mu_i f_i(\theta, x(\theta, p), u) + \nu_i f_j(\theta, x(\theta, p), u) \quad (8)$$

where  $\mu_i, \nu_i$  are the costate variables for firm  $i$ .

The standard conditions for a Nash equilibrium in open-loop strategies are

$$\frac{\partial H_i}{\partial p_i} = 0 \text{ for } i = 1, 2 \quad (9)$$

with

$$\dot{\mu}_i = r\mu_i - \partial H_i / \partial \theta_i, \text{ and } \dot{\nu}_i = r\nu_i - \partial H_i / \partial \theta_j. \quad (10)$$

---

<sup>3</sup>In Chintagunta and Rao's original specification, allowance was made for  $\beta$  and  $\gamma$  to take different values for the two firms whereas the utility from consumption of any good was normalised to unity. I opt for a convention which is closer to the learning literature where  $\beta$  is fixed but utility  $u_i$  varies across the firms.

## 4 A Dynamic Model with Observable Quality

This section is concerned with the case of search goods, where the quality of rival goods is clearly discernible before purchase. As noted above, this corresponds to the “full information” case in the literature on adaptive learning. CR considered only experience goods, but the search goods case provides an interesting point of comparison.

For the moment, let us assume that the qualities of both goods are constant and exogenous and given by  $u_1, u_2$ . Under full information at each point in time, consumers are able to observe the utility from consumption of each good. They are assumed to update their estimates  $\theta$  accordingly. In particular, consider

$$\theta_{it+1} = (1 - \delta)\theta_{it} + \delta u_i \text{ for } i = 1, 2 \quad (11)$$

where  $1 > \delta \geq 0$  is a forgetting parameter. Much of the learning literature has considered learning with no forgetting, for example, fictitious play. However, the assumption of forgetting seems particularly appropriate in this context (indeed CR estimate  $\delta$  at well above 0 from actual sales data). The continuous time equivalent learning rule is

$$\dot{\theta}_i = \delta(u_i - \theta_i) \text{ for } i = 1, 2 \quad (12)$$

Note that here, just as in CR’s original model, price does not affect goodwill. There are arguments for and against this modelling choice. The model is intended to represent the situation of a consumer choosing between products in a supermarket, where prices are clearly displayed. In that sense, because current prices are easily available, the consumer’s choice may not be affected by her knowledge of past prices. On the other hand, a consumer may not check prices again each time he shops. In which case, his choice in a particular period may be determined by his impression about which of the brands is the least expensive, an impression formed by past prices.

In any case, in the present case of observable quality, there is a particularly simple outcome. The Hamiltonian in this case is given by

$$H_i = p_i x_i + \mu_i \delta(u_i - \theta_i) + \nu_i \delta(u_j - \theta_j) \quad (13)$$

Because the updating of goodwill according to (12) does not depend on which good is purchased and hence is not a function of price, the dynamic equilibrium is simply to play a standard duopoly equilibrium at each point in time. That is, the dynamic first order conditions (9) reduce to

$$\frac{\partial H_i}{\partial p_i} = x_i + p_i \gamma x_i x_j = 0 \text{ for } i = 1, 2 \quad (14)$$

or

$$p_i = -\frac{1}{\gamma x_j} \quad (15)$$



Caplin and Nalebuff (1991) show that this first order condition identifies a unique equilibrium in pure strategies to this game. However, it is difficult to obtain an exact solution for the implied price dynamics. However, one has

$$\dot{p}_i - \frac{\dot{x}_j}{\gamma x_j^2} = 0 \quad (16)$$

and one can calculate that at each point in time

$$\theta_i(t) = u_i(1 - e^{-\delta t}) - \theta_i(0)e^{-\delta t}.$$

Or in other words, in the limit,  $\theta_i = u_i$ . That is, the consumer correctly learns the utility she derives from each product. Since in the limit, the sellers cannot influence her opinion, the steady state prices are the same as those arising from a static duopoly (or if the firms were completely myopic).

**Proposition 3** *In the steady state, prices are given by the solution to the static duopoly game where demands are given by the choice rule (5) and  $\theta_i = u_i$ .*

## 5 Experience Goods

In this section, the model is developed under the assumption that the quality of a good is only observable upon consumption. This is CR's original model with slight modifications as noted above. In particular, one can think of the change in goodwill for the consumer following what is now the standard reinforcement learning updating rule (see, for example, Erev and Roth, 1998). That is, if action/good  $i$  is chosen at time  $t$ , then

$$\begin{aligned} \theta_{it+1} &= (1 - \delta)\theta_{it} + u_i \\ \theta_{jt+1} &= (1 - \delta)\theta_{jt} \end{aligned} \quad (17)$$

Moving to the expected motion and continuous time, one obtains

$$\dot{\theta}_i = x_i u_i - \delta \theta_i. \quad (18)$$

The Hamiltonian becomes

$$H_i = p_i x_i + \mu_i (x_i u_i - \delta \theta_i) + \nu_i (x_j u_j - \delta \theta_j). \quad (19)$$

Differentiating  $H_i$  with respect to  $p_i$  and setting to zero, prices must satisfy

$$p_i = -\frac{1}{\gamma x_j} + \nu_i - \mu_i. \quad (20)$$

From this, it is possible to obtain

$$\dot{p}_i - (r + \delta)p_i + \frac{\beta x_i(u_i + u_j)}{\gamma} - \frac{r + \delta}{\gamma x_j} - \frac{\dot{x}_j}{\gamma x_j^2} = 0 \quad (21)$$

As CR calculate, the steady state is characterised by the following equations

$$p_i = \frac{1}{\gamma(r + \delta)} \left( \frac{\beta(u_i + u_j)\delta\theta_i}{u_i} - \frac{(r + \delta)u_j}{\delta\theta_j} \right), \quad x_i u_i = \delta\theta_i \quad (22)$$

First, it is interesting to compare the steady state prices with those obtained in the previous section under full information. From (22) above, one can calculate that

$$p_i = \frac{\beta(u_i + u_j)\delta\theta_i}{u_i\gamma(r + \delta)} - \frac{1}{\gamma x_j}. \quad (23)$$

If  $u_i$  and  $u_j$  are both positive, then the above price will be *lower* than  $-1/\gamma x_j$ , the price obtained under full information (the two prices will coincide if the firms are myopic, i.e.  $r = \infty$ ). Strangely, the additional information available in the full information model, actually makes prices less competitive.

This is not the full story, however. The equations (22) in fact typically are consistent with three distinct steady states. For example, with  $u_1 = u_2 = 1$ , the two firms are in effect identical. However, if one assumes for convenience that  $\delta = 1/2, \gamma = -1, r = 1, \beta = 2$ , there are steady states with  $\theta = (0.645, 1.355)$ ,  $\theta = (1, 1)$ , and  $\theta = (1.355, 0.645)$ , one symmetric, and two highly asymmetric. What is going on?

If one fixes prices and looks at the behaviour of consumer learning in the neighbourhood of a steady state with the same parameter values as before

$$\dot{x}_i = \beta x_i(1 - x_i)(\dot{\theta}_i - \dot{\theta}_j) = \beta x_i(1 - x_i)(2x_i - 1 - \delta(\theta_i - \theta_j)).$$

But from the choice rule (5), with prices fixed,  $\beta(\theta_j - \theta_i) = \log(1 - x_i) - \log x_i$ . This gives

$$\dot{x}_i = \beta x_i(1 - x_i) \left( 2x_i - 1 + \frac{\delta}{\beta} (\log(1 - x_i) - \log x_i) \right). \quad (24)$$

This is a perturbed form of the evolutionary replicator dynamic.<sup>4</sup> One can calculate that it does have a fixed point at  $x_i = 1/2$ , the symmetric outcome one would expect given the firms are themselves symmetric. But one can also calculate that this outcome is unstable. This reflects the result of Rustichini (1999) that under partial information, learning with the exponential choice rule can converge to a suboptimal outcome. For example, the equation

---

<sup>4</sup>For more detailed analysis of the connection between learning processes and the replicator dynamics, see Hopkins (2001).

(24) also has fixed points at  $x_i \approx 0$  and  $x_i \approx 1$ , which will be stable under these learning dynamics.

That is, a consumer can, by pure force of habit, become locked into purchasing one good repeatedly. This is mitigated in practice by the firm which is receiving less custom reducing its price to compensate. For example, from the numerical example above, at the steady state  $\theta = (0.645, 1.355)$ , prices will be  $(0.616, 1.293)$ . This is why in this asymmetric steady state, market shares are  $(0.3225, 0.6775)$  and not  $(0,1)$ . Nonetheless, this model implies a considerable first-mover advantage. If initial conditions are such that the consumer has a marked preference for one firm, one would expect convergence to an asymmetric outcome.<sup>5</sup>

## 6 Experience Goods with an Alternative Learning Model

Sarin and Vahid (1999) propose a new learning model which is particularly applicable when an agent has partial information in the sense of Rustichini (1999). Or, in the present context, it should be adapted to the case of experience goods.

The rule is, if action/good  $i$  is chosen at time  $t$ , then

$$\begin{aligned}\theta_{it+1} &= (1 - \delta)\theta_{it} + \delta u_i \\ \theta_{jt+1} &= \theta_{jt}\end{aligned}\tag{25}$$

The difference between this rule and the rule (17) in the previous section is that now the goodwill toward the good not chosen does not decay. This may seem a slight difference, but it is crucial. Because one's estimate of the quality of the good not chosen will remain above zero, this prevents one becoming locked into the other good through pure choice of habit.

Assuming as before decision rule (5), and moving to expected motion and continuous time, one obtains

$$\dot{\theta}_i = x_i \delta (u_i - \theta_i).\tag{26}$$

It is possible to obtain

$$\dot{p}_i - r p_i + \frac{\delta \beta x_i (u_i - \theta_i + u_j - \theta_j)}{\gamma} - \frac{r}{\gamma x_j} - \frac{\dot{x}_j}{\gamma x_j^2} = 0\tag{27}$$

In the steady state one can calculate from (26) that  $\theta_i = u_i$  for  $i = 1, 2$  and from (27) one can see that prices are  $p_i = -1/(\gamma x_j)$ , just as in (15). Hence, we have

---

<sup>5</sup>Numerical analysis suggests that all three steady states are saddlepoints under the dynamics investigated here. This means that all three are attainable asymptotically.

**Proposition 4** *In the model of experience goods with the Sarin and Vahid learning model, asymptotically the consumer has a correct perception of the quality of the two goods,  $\theta_i = u_i$ , and the two sellers charge the myopic duopoly prices.*

That is, just as in the case of learning under complete information considered in Section 3, there is complete learning. In the limit, the consumer knows the true values of  $u_1$  and  $u_2$ , and the firms cannot influence his beliefs. As a consequence, prices asymptotically approach their myopic levels. The two steady states, under complete information and under partial information with the learning rule of Sarin and Vahid considered in this section, are identical. Of course, this does not imply that the dynamics out of equilibrium are identical. Indeed, the two equations (16) and (27) are quite different. This reflects that here the two firms play the myopic equilibrium only asymptotically, but in the case of full information, prices are set myopically at each point in time.

## 7 Going Beyond a Representative Consumer

The use of a representative consumer makes our models more tractable, but it is not ideal. Equally, assuming that the utility from consumption is constant means that any learning undertaken is trivial. In this section, I investigate the implications of relaxing both assumptions, using a method of aggregation across a population of learners first used in Hopkins (1999).

First, let us assume that there is a large, formally infinite, population of consumers, each with goodwill  $\theta = (\theta_1, \theta_2)$ . Let  $q = \theta_2 - \theta_1$ . Then the beliefs of the population can be described by a distribution  $F(q)$  on  $\mathbb{R}$ . The choice rule (5) can be written

$$x_{1t}(q, p) = \frac{1}{1 + \exp(\beta q_t + \gamma(p_{2t} - p_{1t}))} \quad (28)$$

and demand for the first firm is

$$D_1(F(q), p) = \int x_1(q, p) dF(q). \quad (29)$$

Second, suppose that instead of quality being constant, it is subject to stochastic variation. In particular, assume that  $u_i = \bar{u}_i + \epsilon_i$  for  $i = 1, 2$  and each  $\epsilon_i$  is a random variable, normally distributed with zero expectation and standard deviation  $\sigma$ . So if  $\epsilon = \epsilon_2 - \epsilon_1$ , then  $\epsilon$  is normally distributed with mean zero and standard deviation  $\sqrt{2}\sigma$ .

In this stochastic setting we obtain quite different results.

**Proposition 5** *Under rule (11) and under (25), asymptotically  $F(q)$  converges to a normal distribution  $\Phi(q)$  with mean  $\bar{u}_2 - \bar{u}_1$  and standard deviation  $\sigma\sqrt{2\delta/(2-\delta)}$ .*

That is, the dispersion of the asymptotic distribution of beliefs depends positively on the variance of payoffs and the degree of forgetting. This has important consequences.

**Proposition 6** *Let  $\gamma = -\beta$  so that goodwill is a perfect substitute for price. In the limit,*

$$\lim_{\beta \rightarrow \infty} D_1(F(q), p) = F(p_2 - p_1).$$

Note that in the deterministic settings of Sections 3 and 5, the asymptotic distribution is just a mass point at  $q = \bar{u}_2 - \bar{u}_1$ . Hence, the above result implies as the degree of sensitivity to prices becomes very large, we approach Bertrand competition. For example, if  $\bar{u}_1 = \bar{u}_2$  then the firm which charges a price lower by  $\epsilon$  will have a market share of 1. However, with stochastic utility, because the asymptotic distribution is non-degenerate, price competition is muted. A firm can charge a higher price than its rival and still have a positive market share. For example, assume  $\bar{u}_1 = \bar{u}_2$ , then  $D_1 = \Phi(p_2 - p_1)$ , where  $\Phi$  is the distribution function for the limiting distribution established in Proposition 5, and one can calculate that in a static duopoly equilibrium,

$$p_1 = p_2 = \sqrt{\frac{\delta\pi}{2-\delta}}\sigma > 0.$$

Equilibrium prices are above the competitive level. How far above depends positively on the variance of the noise, and the level of forgetting. If qualities are unequal, then there is no analytic solution. However, numerical analysis reveals that if for example  $\sigma = 1/2$ ,  $\delta = 2/3$ ,  $|\bar{u}_2 - \bar{u}_1| = 1$ , then the disadvantaged firm still has a market share of 15% in equilibrium. Recently, Capra et al. (2000) have examined how a small amount of bounded rationality on the part of sellers can lead to substantial deviation from competitive outcomes under Bertrand competition. Here, learning by consumers fails to be complete allowing substantial softening of price competition.

The case of the model of CR is much more complex. Since the evolution of beliefs depends on which good is purchased, the path of prices will affect what will be the final distribution of beliefs. Simulation for fixed prices suggests that beliefs converge to a bimodal distribution, with accumulations at  $q = -\bar{u}_1$  and  $q = \bar{u}_2$ . This suggests that under this learning rule, the environment will be less competitive still. But this analysis for the moment will have to be left to future research.

## References

Bagwell, K., Ramey, G. (1994). "Advertising and coordination," *Review of Economic Studies* **61**, 153-72

- Börgers, T., Sarin, R.** (2000). “Naive reinforcement learning with endogenous aspirations”, *Int. Economic Review*, **41**, 921-950.
- Camerer, C., Ho, T-H.** (1999). “Experience-weighted attraction learning in normal form games,” *Econometrica*, **67**, 827-874.
- Caplin, A., Nalebuff, B.** (1991). “Aggregation and imperfect competition: on the existence of equilibrium,” *Econometrica*, **59**, 25-59.
- Capra, M., Goeree, J., Gomez, R., Holt C.** (2000) “Learning and Noisy Equilibrium Behavior in an Experimental Study of Imperfect Price Competition,” *Int. Economic Review*, forthcoming.
- Chen, Y., Rosenthal, R.W.** (1996). “Dynamic duopoly with slowly changing customer loyalties,” *Int. J. Industrial Organization*, **14**, 269-296.
- Chintagunta, P.K., Rao, P.V.** (1996). “Pricing strategies in a dynamic duopoly: a differential game model”, *Management Science*, **42**, 1501-14.
- Consumer Union** (1991). “The cola wars”, *Consumer Reports*, **56**, 518-25.
- Erdem T. et al** (1999). “Brand equity, consumer learning and choice”, *Marketing Letters*, **10**, 301-318.
- Erev, I., Barron, G.** (2001). “On adaptation, maximization and reinforcement learning among cognitive strategies”, working paper, Columbia University.
- Erev, I., Haruvy, E.** (2001). “Variable pricing: a customer learning perspective”, working paper.
- Erev, I., Roth, A.E.** (1998). “Predicting how people play games: reinforcement learning in experimental games with unique, mixed strategy equilibria,” *Amer. Econ. Rev.*, **88**, 848-881.
- Fishman, A., Rob, R.** (1998). “Experimentation and competition,” *J. Econ. Theory*, **78**, 299-320.
- Fudenberg, D., Levine D.** (1998). *The Theory of Learning in Games*. Cambridge, MA: MIT Press.
- Heller, D., Sarin, R.** (2001) “Adaptive learning with indirect payoff information,” working paper, University of Chicago.
- Hopkins, E.** (1999). “Learning, matching and aggregation,” *Games Econ. Behav.*, **26**, 79-110.
- Hopkins, E.** (2000). “Two competing models of how people learn in games,” forthcoming *Econometrica*.

- Hopkins, E. and Seymour, R.** (2000). "The stability of price dispersion under seller and consumer learning", forthcoming *Int. Economic Review*.
- Padilla, J.** (1995), "Revisiting dynamic duopoly with consumer switching costs," *J. Economic Theory*, **67**, 520-530.
- Rustichini, A.** (1999). "Optimal properties of stimulus-response learning models," *Games Econ. Behav.*, **29**, 244-73.
- Sarin, R., Vahid** (1999). "Payoff assessments without probabilities: a simple dynamic model of choice", *Games Econ. Behav.*, **28**, 294-309.
- Seetharaman, P.B., Chintagunta, P.K.** (1998). "A model of inertia and variety-seeking with marketing variables," *Intern. J. of Research in Marketing*, **15**, 1-17.
- To, T.** (1996). "Multi-period competition with switching costs: an overlapping generations formulation," *J. Industrial Economics*, **44**, 81-7.