ELSEVIER Journal of Econometrics 89 (1999) 41–56

JOURNAL OF Econometrics

Estimating price expectations in the OTC medicine market: An application of dynamic stochastic discrete choice models to scanner panel data

Füsun F. Gönül*

GSIA, Carnegie Mellon University, Pittsburgh, PA 15213-3890, USA

Abstract

This paper examines the differences in managerial implications of dynamic and reduced-form models when the models are applied to brand choice behavior in the over-the-counter (OTC) back and leg ache medicine market. The models we develop are useful to brand managers and researchers who are interested in dynamic factors that influence choice such as prior experience with brands, future price expectations, and an overall concern for the future. The key findings of our research are that the dynamic model is supported by the data over the single-period model; that past experience with a brand matters in current choice; that consumers' price expectations influence their current choices; and that expectations follow a distinct pattern. © 1999 Elsevier Science S.A. All rights reserved.

JEL classification: M31; C33

Keywords: Dynamic programming; Price expectations; Scanner panel data; Structural model

1. Introduction

Recently, more pharmaceutical manufacturers invest in developing and offering over-the-counter (OTC) equivalents of their prescription (Rx) products[1]. In

*E-mail: fg0k@andrew.cmu.edu

[1] The typical OTC drug is lower in strength than its Rx counterpart, and may also vary in form, size, and flavor.

PII: S0304-4076(98)00054-2

1992 the OTC sales were \$11 billion (Chemical Business, 1992). In 1996 consumers spent an estimated \$16 billion on OTC drugs and the market is expected to grow to \$28 billion by the year 2010. The annual compounded growth rate in the OTC industry is 7.5% according to Drug Topics (1994). As competition intensifies in the OTC pharmaceutical market, various brands of drugs for a particular disease state become available to the consumers without a physician's recommendation. Increased competition results in reduced prices and increased number of choices for the consumers.

Various trade journals report that the competition is intense in areas such as pain relievers, analgesics, allergy, cold, and sinus products, and antacids (Chemical Marketing Reporter, 1995). Most of the time consumers may be reluctant to visit the doctor's office because of its time costs and monetary costs. Consumers may prefer to buy a drug off the shelf (potentially with the advice of the pharmacist) to gain relief from their discomfort. Thus, competition not only lowers prices but also increases accessibility to pharmaceuticals in thousands of retail outlets. Healthcare maintenance organizations (HMOs) support the trend from Rx to OTC because it reduces insurance costs due to fewer physician visits, lab tests, and hospitalizations. The American Medical Association is opposed to the trend because of possible mistreatment of potentially serious illnesses. We model consumers' brand choices in this growing market with state-of-the-art econometric methods. The results of this study should be of interest to the policy makers in the health care environment.

Following the famous Lucas' critique (Lucas, 1976), estimable dynamic stochastic discrete choice models have become increasingly popular in the econometrics literature. According to Lucas, adaptive forecasting techniques, ranging from the simplest exponential smoothing to regression, are good models for short-term forecasting. However, such models are not appropriate for policy evaluations since the underlying preference parameters of the model are not stable.

Since the Lucas' critique, the structural estimation method became popular beginning in labor economics. (The empirical implementation of estimable dynamic stochastic discrete choice models is often referred to as structural estimation, for short.) A good survey of the first structural applications can be found in Eckstein and Wolpin (1989). We accept the Lucas critique as many others and adopt a model that directly enables estimating the parameters of the utility function of the consumers that characterize their individual preferences. We assume that the utility function remains stable over the consumer's lifetime and that the consumer makes optimal decisions in each period with regard to brand choice.

The brand choice literature is large and growing in the marketing literature. Earlier works date to Frank (1962) and Kuehn (1962) who adopt the concept of probability to brand choice phenomena in order to predict market shares. Givon and Horsky (1978) extend the basic Markov and Bernoulli models to include heterogeneity. With the advent of the scanner panels from the grocery stores, household-level brand choice and marketing exposure data have become accessible to marketing researchers in the 1980's (Guadagni and Little, 1983)[2].

Information uncertainty present in the nature of the pharmaceutical products and the absence of physician prescription make switching among OTC drugs difficult for the consumer. For example, if brand A worked in the past for the consumer, then brand B may be risky and costly to try, especially in the absence of free samples that physicians typically dispense for prescription drugs. Television advertising, word of mouth advertising, asking a friend or a relative, asking a physician, a pharmacist, a nurse, or other health professional are options, however, for categories such as minor back and leg ache, the information is usually best obtained by self-experience. Thus, the consumer is mostly self-reliant to find out which brand works for a particular ache. Hence, we expect the prior number of purchases of a brand to be a powerful predictor of future brand choice in the OTC medicine category.

Pharmaceutical manufacturers are often criticized by the Congress and consumer groups regarding their high prices, since consumers shopping for pain medicine are not likely to be as choosy as consumers shopping for other drugstore items (Miller, 1993). While manufacturers argue that their prices are competitive and cover the research and development costs, there is nonetheless an attack on the pharmaceuticals, because the consumers in the medicine market are particularly vulnerable (PhRMA, 1994; U.S. Congress, 1993). Unlike prescription drugs consumers bear the full cost of the OTC pain medicine since health insurance plans do not cover OTC drugs. Thus, we expect price to play an important role in consumers' brand choice decisions.

The proposition that consumers may hold price expectations based on the prices they paid in the past for the same brand has found strong empirical support in the marketing literature. (The reduced-form applications of price expectations include Winer (1986), Helgeson and Beatty (1987), Narasimhan (1989), Kalwani et al. (1990), and Kalwani and Yim (1992). Among the products considered are: durable goods, coffee, liquid laundry detergent, jeans, bicycle, soap and toothpaste.) Further, a variety of hypotheses exist regarding differences in consumer evaluations when prices deviate from expectations in positive and negative directions. (Kalyanaram and Winer, 1995).

We build on these works and formulate that the expected price distribution differs depending on whether the current price is higher or lower than the average price until then. Suppose a consumer is indifferent between brands A and B, and does not expect prices to increase or decrease in the future. Then a lower price should attract more customers to a brand, other things held equal.

[2] Scanner panel data contain longitudinal purchase histories of individuals who volunteer to have their purchases in all stores scanned and stored. More details on the data are in Section 3.1.

On the other hand, if the consumer favors B based on brand image or prior experience with the brand, and expects the price of brand B to be even higher in the next period then the consumer may opt for brand B, even if the current price of B is relatively higher than A. Thus, a positive impact of current price on brand choice is possible in a dynamic framework. We estimate the probability distribution of expected future prices based on purchase data.

In summary, we attempt to improve the existing brand choice models on two grounds: (1) we adopt a structural estimation method to estimate the stable preference parameters of the consumers and (2) in this structural setting we estimate the pattern of consumers' price expectations in the future. The role of price expectations on current purchases is certainly not new to the marketing literature (see above for references). However, we are the first to incorporate them in a dynamic, estimable, infinite-horizon framework.

The rest of the paper is arranged as follows. Section 2 presents the model, Section 3 describes the data and estimation, Section 4 discusses the results, and Section 5 concludes.

2. Model

2.1. Dynamic Programming Model

At any time $t = 1, \ldots, \infty$ consumers maximize the objective function

$$\mathrm{E}\left\{\sum_{\tau=t}^{\infty} \delta^{\tau-t} \sum_{j=1}^{J} U_j(\tau) d_j(\tau) | S(t)\right\}, \tag{1}$$

where j is the brand index ($j = 1, \ldots, J$), $U(\cdot)$is the per-period utility, $d_j(t)$ are the binary choice variables, $\delta > 0$ is the monthly discount factor, and $S(t)$ denotes the state space at time t. The state space consists of all the information available to the individual at time t such as prices and past choices.

Maximization of Eq. (1) is accomplished by the choice of the optimal sequence of control variables $\{d_j(t)\}$ for $t = 1, \ldots, \infty$. Only one brand is chosen by a consumer at a given time, that is, the control variables satisfy the condition that $\sum_{j=1}^{J} d_j(t) = 1$. The maximization is achieved by the *principle of optimality* that states that, whatever the initial state and decision, the remaining decisions must be optimal with regard to the state resulting from the first decision (Bellman, 1957). We define the maximal expected value of the discounted lifetime utility as

$$V(S(t)) = \max \mathrm{E}\left\{\sum_{\tau=t}^{\infty} \delta^{\tau-t} \sum_{j=1}^{J} U_j(\tau) d_j(\tau) | S(t)\right\}, \tag{2}$$

where the maximization is over $\{d_j(t)\}$. The value function $V(S(t))$ depends on the state space at t and can be written as $V(S(t)) = \max\{V_j(S(t))\}$ where the maximization is over $j = 1, \ldots, J$. The brand-specific value functions satisfy the Bellman equation for all t:

$$V_j(S(t)) = U_j(S(t)) + \delta \mathrm{E}\{V(S(t+1)) | S(t), d_j(t) = 1\}. \tag{3}$$

The brand-specific value function explicitly links the choice at time t to the optimal time path from $t+1$ onwards. The formulation also indicates that future choices are made optimally for any given current period decision.

The consumer is assumed to know the distribution of future prices. As discussed ahead in Section 2.4 we assume a discrete three-point distribution for prices. Hence, the expectation term becomes the sum of the products of future value functions conditional on price points and their respective probabilities.

2.2. Empirical issues

The per-period utility function from choosing brand j at time t is

$$\begin{aligned} U_j(t) = \alpha_j &+ \beta_{Nj}(Num\_Pur)_{jt} + \beta_{Nj-Squared}(Num\_Pur)^2_{jt} \\ &+ \beta_{Pj} Price_{jt} + \varepsilon_{jt}, \end{aligned} \tag{4}$$

where α_j are brand-specific constants (one of which is normalized to zero for identification of others), $(Num\_Pur)_{jt}$ is the number of previous purchases, and $Price_{jt}$ is the price of brand j at time t. To capture potential diminishing returns in the prior number of purchases we enter the quadratic formulation for the $(Num\_Pur)_{jt}$ variable[3].

Possible addiction to the brand or a relatively high level of satisfaction with the brand increases the linear coefficient of the number of purchases, and hence makes a switch more difficult. If the quadratic coefficient of the number of purchases is positive, then switching becomes even more difficult as the number of prior purchases increases. If the quadratic coefficient is negative, diminishing returns set in after a certain number of prior purchases of the same brand.

In a purely static framework the price coefficients are expected to be negative consistent with the inverse demand curve. However, in a dynamic framework we may have positive coefficients as discussed in the introduction. If a relatively high price causes anticipation of even higher prices in the future then it may induce a purchase in the present period, since the consumer maximizes lifetime utility and not just current utility.

---

[3] We thank an anonymous referee for this suggestion.

The consumer knows the distribution of the random term (ε_{jt}) and makes a draw from this distribution in each period. If the random terms are distributed extreme value i.i.d. then the expected value function that appears in Eq. (3) has the following closed-form expression:

$$\mathrm{E}\{V(S(t))\} = \gamma + \ln\left[\sum_{j'=1}^{J} \overline{\exp(V_{j'}(S(t)))}\right], \tag{5}$$

where γ is Euler's constant (Rust, 1987). The upper bar indicates the part of the value function that is deterministic.

In each time period the consumer makes a selection that maximizes the lifetime utility function. The modeler observes the choice and estimates the choice probabilities. The choice probabilities achieve the multinomial logit formula due to the extreme value assumption on the error terms

$$\Pr ob(d_j(t) = 1|S(t)) = \frac{\exp(\overline{V_j(S(t))})}{\sum_{j'=1}^{J}\exp(\overline{V_{j'}(S(t))})}. \tag{6}$$

The choice probability for each brand is influenced by the prices and past purchases of *all brands* because of the term in the denominator.

The medicine purchase problem is best cast in the *infinite horizon* because the nature of the product category does not lend credibility to the finite horizon assumption, unlike, for example, baby diapers where the horizon is two or three years. The solution of the infinite horizon problem rests on a theorem that states that under certain conditions, the decision process reaches stationarity, and there exists an expected value function such that $EV(S(t)) = EV(S(t+1))$ (Bertsekas, 1976)[4].

We maximize the log-likelihood function that consists of the sum of the logarithms of the choice probabilities in Eq. (6). The maximum likelihood estimation routine envelops the dynamic programming routine and iterates over the parameter values. The dynamic programming problem is solved in order to satisfy the infinite horizon stationarity condition within each iteration over the parameter space.

We conduct the maximization using the OPTMUM library of GAUSS, on a personal computer with a speed of 60 MHz. For a given set of parameters and an arbitrary starting point for the expected value function, we initially set the

[4] The consumers in the data set have been in the market for a while according to a questionnaire they fill out prior to being admitted to the survey panel. Hence, we can assume that the choice processes of the consumers have reached stationarity long before the observation interval began that suits the infinite horizon assumption.

length of the horizon to two periods. We solve the dynamic programming problem for all values of the elements of the state space using the backward recursion method. We increment the number of time periods and evaluate the expected value functions until their difference in consecutive time periods become smaller than the tolerance limit of the optimization algorithm. If convergence occurs such that $EV(S(T)) = EV(S(T+1))$ we stop the backward recursion algorithm and use the converged expected value functions to evaluate the likelihood function[5]. We repeat the entire process for a new set of parameter values and continue until the likelihood function converges.

2.3. State space

At any given time period the consumer uses the available information on brands in making a selection. The information set is summarized in the state space that includes past choices and distribution of future prices ($f(\cdot)$) given current prices:

$$S(t)\colon\{(Num\_Pur)_{jt} = (Num\_Pur)_{j,t-1} + d_j(t-1); f(Price_{j,t+1}|Price_{jt});\ \forall j\}, \tag{7}$$

where $(Num\_Pur)_{j0}$ is obtained from the initialization period in the data.

Computational issues: Ideally, the dynamic programming problem has to be solved for all possible *future* price levels of the 4 brands. However, since price is a continuous variable it is prohibitively expensive to consider all possible prices. For example, if the number of possible discrete values of $(Num\_Pur)_{jt}$ and $Price_{jt}$ is both five then the size of the state space becomes $5^4 \times 5^4 = 390{,}625$. And if stationarity is reached within 40 evaluations to satisfy the infinite horizon condition, then the number of times the dynamic programming problem has to be solved exceeds 15 million for a given set of parameter values! To contain the computational costs, we limit the possible levels of expected price to three: one standard deviation higher than average price, average price, and one standard deviation lower than average price. Note that categorizing prices to three levels affects only the future component of the valuation functions. Current period prices are continuous and used as their observed values in Eq. (4).

2.4. Price expectations

Unrestricted Model: We form and test the hypothesis that consumers react differently if the current price is higher than the historical average than if the

---

[5] With this method we are able to reach stationarity after 40–50 periods within an iteration.

current price is lower than the historical average. Then expected price in the next period becomes

$$\begin{aligned}
&\mathrm{E}(Price_{j,t+1}|Price_{jt} > \overline{Price_{jt}}) \\
&\quad = q_{11}\{\overline{Price_{jt}} + (Std.Dev.)_{jt}\} + q_{12}\overline{Price_{jt}} + q_{13}\{\overline{Price_{jt}} - (Std.Dev.)_{jt}\} \\
&\quad = \overline{Price_{jt}} + (q_{11} - q_{13})(Std.Dev.)_{jt}, \\
&\mathrm{E}(Price_{j,t+1}|Price_{jt} \leqslant \overline{Price_{jt}}) \\
&\quad = q_{01}\{\overline{Price_{jt}} + (Std.Dev.)_{jt}\} + q_{02}\overline{Price_{jt}} + q_{03}\{\overline{Price_{jt}} - (Std.Dev.)_{jt}\} \\
&\quad = \overline{Price_{jt}} + (q_{01} - q_{03})(Std.Dev.)_{jt},
\end{aligned} \tag{8}$$

where$\sum_{k=1}^{3} q_{1k} = 1$, $q_{1k} \geqslant 0$, and $\sum_{k=1}^{3} q_{0k} = 1$, $q_{0k} \geqslant\ = 0$. These are probabilities and are estimated jointly with the rest of the model parameters. In the data, the mean and standard deviation of the prices $(\overline{Price_{jt}}, (Std.Dev.)_{jt})$ change with each purchase as consumers update their purchases over time. Note that we do not simply plug in expected prices in the future part of the Bellman equation but rather sum over the future price distribution conditional on current price as discussed in Section 2.1.

Restricted Model (1): At the outset it is informative to test the main hypothesis depicted by the unrestricted case that price expectations differ depending on the value of current price relative to the historical average. For example, if $q_{0k} = \frac{1}{3} = q_{1k}$ for $k = 1, 2, 3$ then expectations are no longer conditional on current prices but become $\mathrm{E}(Price_{j,t+1}) = \overline{Price_{jt}}$. The same model can be obtained by imposing the less severe restrictions $q_{11} - q_{13} = q_{01} - q_{03} = 0$. In this model, lower prices should attract more customers since price expectations are neutral with respect to the current price[6].

Restricted Model (2): A relatively high current price may be attractive to consumers and induce them to buy now if they expect the price of the brand to be even higher in the next period. Conversely, if the current price is relatively low and if it is expected to fall even more in the next period then consumers may prefer to wait. Both scenarios explain the possibility of positive coefficients on prices. This is achieved by $q_{11} = q_{03} = 1$ (which imply $q_{12} = q_{13} = 0$ and

[6] A more general statement for Restricted Model (1) is given by the restriction $q_{11} - q_{13} = q_{01} - q_{03}$ so that

$$\mathrm{E}(Price_{j,t+1}) = \overline{Price_{jt}} + (Std.Dev.)_{jt}(q_{11} - q_{13}).$$

$q_{01} = q_{02} = 0$) or by the less extreme case, $q_{11} > \frac{2}{3}$ and $q_{03} > \frac{2}{3}$. Then higher than average prices generate disproportionate pessimism and higher prices are expected to continue in the future. Lower than average prices generate disproportionate optimism and lower prices are expected in the future period. For example, in the extreme case

$$E(Price_{j,t+1} | Price_{jt} > \overline{Price_{jt}}) = \{\overline{Price_{jt}} + (Std.Dev.)_{jt}\},$$
$$E(Price_{j,t+1} | Price_{jt} \leqslant \overline{Price_{jt}}) = \{\overline{Price_{jt}} - (Std.Dev.)_{jt}\}. \tag{9}$$

This leads to the interesting result that the price effect may be positive in contrast to Restricted Model (1).

Other Restricted Models: The expected direction of prices can be in the opposite direction of current prices if $q_{11} = q_{12} = 0$ and $q_{02} = q_{03} = 0$ (which imply $q_{13} = q_{01} = 1$). In other cases expectations can be only inflationary ($q_{11} = q_{01} = 1$) or only deflationary ($q_{13} = q_{03} = 1$). We test these restrictions by using the standard errors of the estimates of the $\boldsymbol{q}_1$, $\boldsymbol{q}_0$ parameters from the unrestricted model.

Reduced-Form (*Single-Period*) *Model*: The models above collapse to a single-period discrete choice model when the discount factor is zero ($\delta = 0$). The single-period model does not allow for a price expectation structure in determining choice probabilities, by construction.

3. Data and estimation

3.1. Data

The scanner panel data is collected from various cities in the U.S. by A.C. Nielsen Company. The sample spans a two-year period where daily purchase information is available on 332 buyers of back and leg ache medicine. We allow for an initialization period of 6 months to initialize the previous experience with brands and average prices. The total number of purchase observations is 1205.

The market is dominated by four brands, where brand 1 enjoys 45% market share and has the lowest unit price (18.73¢ per tablet, with a standard deviation of 2.85¢). Brand 2 is the second leading brand with 22% share and the highest unit price (25.36¢ with std. dev. 2.11¢). The third and fourth brands each have about 17% market share and are in the medium price range (22.86¢ with std. dev. of 1.29¢ and 21.73¢ with std. dev. of 2.28¢). We leave out the last third of the observation interval for each household as holdout period in order to make prediction comparisons across various models.

The average number of purchases is 8.7 with a minimum of 2 and a standard deviation of 12. The time is measured in months. The category is characterized by a high repeat purchase tendency. For example, during the two-year observation period 36% of the consumers are observed to always have bought the same brand.

3.2. Comparison of models

Likelihood ratio tests show that the Restricted Model (1) is rejected in favor of the unrestricted dynamic model. This is verified by both the likelihood ratio test ($\chi^2_{(2)} = 142, p < 0.01$)[7] and the asymptotic normal tests (both $q_{11} - q_{13}$ and $q_{01} - q_{03}$ are significantly different than zero, $p < 0.01$). The equality $q_{11} - q_{13} = q_{01} - q_{03}$ that yields a more general case of the Restricted Model (1) is also rejected, $p < 0.01$. Similar asymptotic normal tests show that Restricted Model (2) is also rejected. Other restricted versions mentioned in Section 2 are likewise rejected.

The AIC, BIC values, and hit rates in the holdout period in Table 1 also show that the unrestricted dynamic model outperforms restricted dynamic models and single-period model.

We also estimate Bernoulli and Markov models adapted from the Linear Learning Model of Givon and Horsky (1978)[8]. The choice probability is

$$\Pr ob(d_j(t) = 1) = \frac{\exp(Y_j(t))}{\sum_{j'=1}^{J} \exp(Y_{j'}(t))},$$

where $Y_j(t) = \alpha_j + \beta_j \, Repeat_{jt} + \lambda_j \ln(\Pr ob(d_j(t-1)), Repeat_{jt} = 1$ if brand j has been bought in the previous purchase occasion and 0 otherwise. The estimation results are such that $\log - L = -452.25$, $AIC = 926.51$, $BIC = 979.00$, and the holdout hit rate is 89.76%. Although the Markov–Bernoulli model performs better than the unrestricted dynamic model in the holdout period, it achieves a lower hit rate in the estimation period (76.29% < 81.67%).

The verdict on the performance of our approach as a forecasting tool is still undecided since this is an analysis based on a single data set. However, the structural estimation method we use is more appropriate to make policy evaluations than reduced-form models. As we state in the Introduction the predictive performance of reduced-form models are typically satisfactory. In fact

[7] The chi-squared statistic is twice the log-likelihood difference with the degrees of freedom equal to the number of restrictions in the restricted model. The restrictions in this case are: $q_{11} - q_{13} = 0$ and $q_{01} - q_{03} = 0$.

[8] We thank an anonymous referee for this suggestion.

Table 1
Parameter estimates and asymptotic normal statistics

Variables	Unrestricted dynamic model	Restricted dynamic model	Single-period model
Brand-specific intercepts			
α_1	− 0.520 ***	− 0.202 ***	5.986 ***
	(− 10.434)	(− 2.479)	(130.952)
α_2	− 10.620 ***	− 10.717 ***	0.097 ***
	(− 199.275)	(− 219.322)	(2.122)
α_3	1.077 ***	1.068 ***	12.295 ***
	(22.387)	(22.091)	(268.881)
Previous number of purchases (linear and quadratic)			
β_{N1}	5.520 ***	5.482 ***	1.950 ***
	(90.110)	(111.205)	(42.638)
β_{N1}-Squared	− 0.137	− 0.125 *	− 0.219 ***
	(− 1.725)	(− 1.935)	(− 4.775)
β_{N2}	4.217 ***	4.157 ***	0.869 ***
	(78.265)	(66.284)	(18.914)
β_{N2}-Squared	0.638 ***	0.074	− 0.026
	(9.916)	(0.845)	(− 0.564)
β_{N3}	7.457 ***	7.599 ***	3.101 ***
	(143.706)	(144.592)	(67.812)
β_{N3}-Squared	0.248 ***	0.226 ***	− 0.064
	(18.699)	(4.639)	(− 1.391)
β_{N4}	2.941 ***	2.846 ***	2.509 ***
	(55.490)	(46.623)	(54.861)
β_{N4}-Squared	0.068	0.034	− 0.193 ***
	(1.247)	(0.398)	(− 4.218)
Price			
β_{P1}	− 0.085 ***	− 0.079 ***	− 0.029
	(− 3.247)	(− 6.654)	(− 0.744)
β_{P2}	0.382 ***	0.414	0.178 ***
	(30.128)	(0.014)	(7.096)
β_{P3}	− 0.327 ***	0.359 ***	− 0.362 ***
	(− 66.698)	(− 36.559)	(− 7.913)
β_{P4}	0.110***	0.096	0.195 ***
	(9.249)	(0.124)	(4.274)
Price expectation probabilities			
q_{11}	0.606 ***	0.333	
	(17.292)	(Fixed)	Not applicable
q_{12}	0.265 ***	0.333	
	(24.339)	(Fixed)	Not applicable

Table 1 (*continued*)

Variables	Unrestricted dynamic model	Restricted dynamic model	Single-period model
q_{13}	0.129 ***	0.333	
	(12.187)	(Fixed)	Not applicable
q_{01}	0.098 ***	0.333	
	(11.563)	(Fixed)	Not applicable
q_{02}	0.226 ***	0.333	
	(20.991)	(Fixed)	Not applicable
q_{03}	0.677 ***	0.333	
	(18.368)	(Fixed)	Not applicable
Log-Likelihood	− 357.541	− 428.743	− 477.735
AIC	753.082	887.486	985.470
BIC	843.749	959.065	1057.049
Hit Rate in the Holdout Period	82.23%	81.02%	80.41%

Notes: Asymptotic normal statistics are placed in parentheses. A sign of (***) denotes significance at 0.001 level, (**) significance at 0.05 level, and (*) significance at 0.01 level. The intercept for the fourth brand is normalized to zero.

forecasting is the main reason for estimating inexpensive reduced-form models and updating them as current data become available. However, such models do not allow policy evaluations based on stable preference parameters.

4. Discussion

It appears that for some brands the state dependence increases at an increasing rate (see the *number of purchases* coefficients). As the number of prior purchases for brands 2 and 3 increases, it becomes increasingly difficult to shift from these brands since the quadratic effects are also positive and significant. Brands 1 and 4 cause inertia as well, although to a lesser extent since their quadratic coefficients are insignificant. Because of the information uncertainty in the nature of the product category, consumers appear to be reluctant to switch brands once they find a medicine that works. The relative ease of switching differs across brands, however. While brand 3 is the hardest to shift from, brand 4 is the easiest to shift. We do not have enough knowledge on the chemical ingredients of brand 3 to speculate whether drug addiction can be more of a factor for this brand vis-à-vis other brands.

All probabilities related to *price expectations* are significantly different from zero and one, thus supporting the three-point expected price distribution. As discussed in Section 3.2 the Restricted Model (1) that does not allow expectations to differ by the current price level is rejected. Thus consumers appear to differ in their expectations about OTC medicine prices depending on the current price level.

Restricted Model (2) is also rejected as discussed in Section 3.2. That is, consumers do not expect prices to continue rising once they are above the historical average, nor do they expect them to continue falling once they are below the historical average. However, since we find that $\hat{q}_{11}$ is the highest probability (61%) and $\hat{q}_{03}$ is the highest probability (68%), the unrestricted dynamic model resembles Restricted Model (2), although it is not as extreme.

For example, when the current price is higher than the average, consumers expect the price to be strictly higher about 61% of the time in the next period. They expect it to fall strictly below average about 13% of the time. Hence, consumers become pessimistic in the presence of current high prices. In contrast, if the current price is lower than the average, consumers expect the future price to be lower than the average by about $\frac{2}{3}$ probability and are pessimistic only with about 10% probability. Thus, current prices strongly dominate expectations and consumers are flexible, not sluggish, in updating their expectations.

Brand 1 has the lowest *price* and the highest market share. Since the effect of previous purchases are not especially high for this brand, low price appears to be driving the market share for this brand. Although the brand has a negative and significant price coefficient in both of the dynamic models, the single-period model assigns an insignificant price coefficient to the leading brand.

The single-period model also appears to suggest that consumers buy brands 2 and 4 even when price increases. The seemingly wrong result is corrected by the dynamic model even in the restricted version with flat expectations. When prices are not expected to change in the future as in the Restricted Model (1), consumers are drawn more to a brand only if the price is lower. (In this model positive price coefficients are found to be insignificant.)

The unrestricted model yields positive price coefficients for brands 2 and 4, however, offers an explanation based on dynamic orientation of the consumers and their price expectations. When the current price is higher than the historical average for a brand, a strictly higher price is expected with about 61% probability and an average or higher than average price is expected with about 87% probability. Hence, price is expected to fall only about 13% of the time if the current price is relatively high. This would make the current high price appear attractive and induce purchases.

Conversely, if the current price is lower than the historical average for the brand, a strictly lower price is expected with a relatively high probability (68%) and an average or lower than average price is expected with 90% probability. Hence, only 10% of the time consumers expect the price to increase if the current price is relatively low. Then consumers may prefer to defer the purchase to the next period. Thus, the dynamic model offers an empirical explanation for positive coefficients on prices, while the reduced-form (single-period) model appears to yield coefficients with wrong signs.

5. Conclusion

We have estimated an infinite horizon brand choice problem for over-the-counter medicine purchases. The results imply that consumers (i) take into account the discounted future utility while making current decisions, and (ii) hold a certain pattern of price expectations when they make current brand choices. Consumers expect prices to rise further in the future if prices are relatively high today compared to the historical average. Their pessimism about future prices is mitigated, if current prices are relatively low.

The main limitation of our model is computational, a curse common to all dynamic programming models[9]. Conceptually, it is possible to extend the state space to accommodate other consumer-specific and store-specific variables. However, each addition multiplies the size of the state space. For example, if household size is added to the model then the dynamic programming problem has to be solved for each value of the household size, given a set of parameter values. The solution would have to be repeated for a new set of parameter values within a maximum likelihood iteration. Hence, the number of times the dynamic programming problem has to be solved grows exponentially with each state variable. Since the computations are lengthy (over two hours for an iteration of the unrestricted dynamic model) we economize on the number of household-specific variables.

Computational cost hinders us from including habit persistence, serial correlation in prices and other unobserved effects that might be correlated with prices over time. Other extensions may be to allow for the price expectations to follow a second-order Markov structure; or incorporating the impact of more drastic price changes by conditioning on price levels other than the historical average. Although it is analytically possible to introduce relatively more complex expectations, computational cost hinders us.

It is conceivable that consumers may consider the future to varying degrees. Some consumers may discount the next n periods heavily while some consumers may not discount the future at all. A way to tackle the heterogeneity in future orientation is by assuming that the discount factor is a random variable. However, attempting to empirically estimate the probability distribution of the discount factor is problematic in the dynamic programming context. To the best of our knowledge no dynamic programming researcher has estimated the discount factor as a parameter let alone assumed it random. This is because the solution of the optimization algorithm is critically dependent on the discount factor, and varying the discount factor during maximum likelihood estimation

[9] It is imperative for empirical modelers to reduce the state space to manageable dimensions in order to cope with the curse of dimensionality, a phrase attributed to R. Bellman.

(as if it is an estimable parameter) results in convergence problems (see, for example, Rust, 1987; Eckstein and Wolpin, 1989; Hotz and Miller, 1993).

We have not modeled the timing of purchases or equivalently, the possibility of no-choice as an alternative. This is because we are focusing on brand choice whenever a purchase occurs. Modeling the timing of purchase (or the no-choice alternative) would imply extending the state space to consider inventory as well. In Gönül and Srinivasan (1996) a purchase/nonpurchase model with fixed prices is estimated in a dynamic programming context. Since the state space is already large in this model, we abstain from inventory decisions and focus on brand choice only. In addition, we have not modeled the firm's side formally. This was achieved in Gönül and Shi (1998) where both the cataloguer firm and the consumer maximize their objectives simultaneously over the time horizon.

Adding the aforementioned factors could improve the hit rates and prediction of market shares if prediction is a key issue. However, as we state in the introduction our goal is not prediction or obtaining the best fit to the data. Perhaps the best forecast is achieved by exponential smoothing and other inexpensive reduced-form methods. Our goal is to develop a model with stable parameters that is useful for policy evaluation. For this purpose we develop a choice model and assume that consumers behave the way we propose in the infinite horizon. Hence ours is an as if model. We estimate the consumers' preference (utility) parameters having imposed a particular structure on the consumer purchase data. All we can conclude is that the data do not reject that consumers behave in the manner governed by the stochastic model we propose.

While there is a growing body of literature on health care costs, the studies have been limited to hospital and physician fees, growth of managed care organizations, and insurance coverage for prescription products. In this paper we explored consumer buying behavior when shopping for nonprescription products. We examined the impact of brand switching costs, inferred from repetition of past purchases, price sensitivities, and future price expectations on current brand choice, assuming consumers maximize lifetime utility subject to uncertainties in the market. The issues discussed here are helpful to policy makers in managed care organizations, pharmaceutical companies, physician associations, as well as the federal government and consumer advocates.

References

Bellman, R., 1957. Dynamic Programming. Princeton University Press, Princeton.

Bertsekas, D., 1976. Dynamic Programming and Stochastic Control. Academic Press, New York.

Chemical Business, 1992. Rx Drugs Invade the OTC Market, vol. 14 (11) pp. 8–10.

Chemical Marketing Reporter, 1995. Pharmaceuticals '95: OTCs – a remedy for growth, vol. 248, no. 12 (September 18), SR12-SR14.

Drug Topics, 1994. Oral OTC Market will Sparkle Rx-to-OTC Switches, vol. 138, no. 9 (May 9), pp. 37-40.

Eckstein, Z., Wolpin, K.I., 1989. The specification and estimation of dynamic stochastic discrete choice models. Journal of Human Resources 24, 562–598.

Frank, R.E., 1962. Brand choice as a probability process. Journal of Business 35, 43–56.

Givon, M., Horsky, D., 1978. Market share models as approximators of aggregated heterogeneous brand choice behavior. Management Science 24 (13), 1404–1416.

Gönül, F.F., Shi, M., 1998. Optimal mailing of catalogs: A new methodology using estimable structural dynamic programming models. Management Science forthcoming.

Gönül, F.F., Srinivasan, K., 1996. Estimating the impact of consumer expectations of coupons on purchase behavior : a dynamic structural model. Marketing Science 15 (3), 262–279.

Guadagni, P.M., Little, J.D.C., 1983. A logit model of brand choice calibrated on scanner data. Marketing Science 2 (3), 203–238.

Helgeson, J.G., Beatty, S.E., 1987. Price expectation and price recall error : an empirical study. Journal of Consumer Research 14 (3), 379–386.

Hotz, V.J., Miller, R.A., 1993. Conditional choice probabilities and the estimation of dynamic models. Review of Economic Studies 60, 497–529.

Kalwani, M.U., Yim, C.K., Rinne, H.J., 1990. A price expectations model of customer brand choice. Journal of Marketing Research 27, 251–262.

Kalwani, M.U., Yim, C.K., 1992. Consumer price and promotion expectations: an experimental study. Journal of Marketing Research 29, 90–100.

Kalyanaram, G., Winer, R.S., 1995. Empirical generalizations from reference price research. Marketing Science: Special Issue on Empirical Generalizations in Marketing 14 (3), G161–G169.

Kuehn, A.A., 1962. Consumer brand choice as a learning process. Journal of Advertising Research 2, 10–17.

Lucas, R.E., 1976. Econometric policy evaluation: a critique. In: Brunner, K., Meltzer, A., (Eds.) The Phillips Curve and Labor Markets, vol. 1 of Carnegie-Rochester Conference Series on Public Policy, North-Holland, Amsterdam, pp. 19–46.

Miller, C., 1993. Drug marketers 'On the alert'. Marketing News 27 (1), 1.

Narasimhan, C., 1989. Incorporating consumer price expectations in diffusion models. Marketing Science 8 (4), 343–357.

PhRMA, 1994. Pharmaceutical marketing practices: questions and answers. PhRMA Marketing Practices Committee, September.

Rust, J., 1987. Optimal replacement of GMC bus engines: an empirical model of harold zurcher. Econometrica 55, 999–1035.

Congress, Office of Technology Assessment, 1993. Pharmaceutical R and D: Costs, Risks, and Rewards. OTA-H-522. U.S. Government Printing Office, Washington, DC.

Winer, R.S., 1986. A reference price model of brand choice for frequently purchased products. Journal of Consumer Research 13, 250–256.