WORKING PAPER
ALFRED P. SLOAN SCHOOL OF MANAGEMENT

ESTIMATING THE PROPORTION OF "ALWAYS BUY" AND "NEVER BUY" CONSUMERS: A LIKELIHOOD RATIO TEST WITH SAMPLE SIZE IMPLICATIONS

Manohar U. Kalwani, M.I.T.
Donald G. Morrison, Columbua University

WP 863-76  JUL 26 1976  July 1976

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
50 MEMORIAL DRIVE
CAMBRIDGE, MASSACHUSETTS 02139
ESTIMATING THE PROPORTION OF "ALWAYS BUY" AND "NEVER BUY" CONSUMERS:
A LIKELIHOOD RATIO TEST WITH SAMPLE SIZE IMPLICATIONS

Manohar U. Kalwani, M.I.T.
Donald G. Morrison, Columbia University

WP 863-76
JUL 26 1976
DEWEY LIBRARY

July 1976
ABSTRACT

The likelihood ratio test is presented as a natural method to test for the presence of "always buy" and "never buy" consumers. The purchase sequence lengths and sample sizes required to estimate the proportions of these buyers are determined through the use of simulated data. Our method was then applied to two sets of well known data first analyzed by Aaker [1] and Montgomery [8]. The empirical findings show that although over half of the consumers never bought a given brand, the specific addition of a spike at p=0 to provide for "never buy" consumers did not form a superior fit over the no-spike model.
INTRODUCTION

Many stochastic models of buyer behavior assume a zero-order process. In each of these models, consumers are assumed to have different but constant over time purchase probabilities. As a result of this, a zero-order model is completely specified by the distribution of purchase probabilities across the population of consumers. The beta distribution, due to its flexibility and tractability has been frequently used as a functional form to characterize this heterogeneity in purchase probability across consumers. The parameters of the purchase probability distribution (e.g., two parameters of the beta distribution) can be estimated from panel data which contain brand purchase information of many households over time.

Often the panel contains many households who have never purchased a given brand and some households who have bought the brand on every purchase occasion. Do such data imply that the purchase probability distribution has mass points at its extremities of p=0 and/or p=1 (p denotes brand purchase probability)? These mass points (or spikes) would represent the proportion of "always buy" and "never buy" consumers. We suggest the likelihood ratio test as the natural method to statistically test for and estimate the size of these spikes.

The major contribution of this paper is to determine the purchase sequence lengths and sample sizes required to use the likelihood ratio test to detect and estimate spikes. Virtually all of the published data use purchase sequence lengths of five or less to estimate the parameters of the purchase probability distribution, but even with very large sample sizes this length is simply not adequate to test for spikes. In order to illustrate
our methodology the likelihood ratio test is used to estimate the spikes for two sets of empirical data that have been used in published research work in marketing, namely the Aaker data [1] for a frequently bought consumer good and Montgomery's [8], [9] Crest toothpaste data.

The paper is organized as follows: We first describe the three alternative models with zero, one or two spikes. We then present the likelihood function approach to estimate the model parameters. This is followed by the use of the likelihood ratio test to detect and estimate spikes. Purchase sequence length and sample size requirements are then determined through the use of simulated data. The application of the likelihood ratio test to the empirical data is presented next. The end results are some "rules of thumb" for sample sizes and purchase sequences lengths needed in estimating the parameters of zero order models.

THE THREE MODELS

Let \( f(p) \) denote a specific functional form of the purchase probability distribution for a particular brand (say, Brand 1). Each consumer has some probability, \( p \), of purchasing Brand 1 which takes a value between zero and one; \( f(p) \) denotes the density of consumers who have that purchase probability. For ease of reference and clearer understanding let us introduce and define the following symbols:

- \( N = \) Number of consumers in the product class.
- \( n_j = \) Number of consumers who make \( j \) purchases of Brand 1 and \((k-j)\) purchases of Brand 0 ('all other' aggregate class) given \( k \) purchase occasions.
- \( k = \) Length of purchase string, that is, the total number of occasions for purchasing Brand 1 or 0.
- \( p_j = k^j R_j = \) Expected probability of making \( j \) purchases of Brand 1 out of a total of \( k \) purchases.
\[ c_j^k = \text{Number of ways in which } j \text{ objects can be chosen out of } k \text{ objects.} \]

\[ f_b(j|p,k) = \text{Binomial probability of making } j \text{ purchases on } k \text{ trials given purchase probability is } p. \]

\[ \alpha = \text{Proportion of consumers who 'never buy' Brand 1.} \]

\[ \beta = \text{Proportion of consumers who 'always buy' Brand 1.} \]

\[ j = 0, 1, 2 \ldots k. \]

\[ \bar{j} = \text{Average purchase frequency.} \]

Table 1 lists in summary form the three alternative models tested in this paper. Along with a pictorial presentation, Table 1 contains a brief description, an expression for the probability \( p_j \), and average purchase frequency for the three models.

**PARAMETER ESTIMATION**

We require the maximum value of the likelihood function for each alternative model to perform the likelihood ratio test. The likelihood function for a model \( M \) (where \( M=0,1,2 \)) given purchase frequency data, \( n_j \)'s, can be written as:

\[
\ell(M) = \ell(n_0, n_1, \ldots n_k; \theta) = \frac{N!}{n_1! n_2! \ldots n_k!} (p_0)^{n_0} (p_1)^{n_1} \ldots (p_k)^{n_k}, \tag{1}
\]

where \( M = \text{model identification number,} \)

\( \theta = \text{parameters of the model, } M. \)

In equation (1) the \( p_j \)'s are, of course, functions of the parameter vector \( \theta \) (in the unconstrained case of Model 2, for instance, \( \theta \) represents \( m, n, \alpha \) and \( \beta \)) as shown in Table 1. The maximum likelihood estimates can be obtained by maximizing \( \ell(M) \) above with respect to \( \theta \). Since maximizing
<table>
<thead>
<tr>
<th>Figure</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Graph 1" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Brief Description</th>
<th>Model 0 (No Spikes)</th>
<th>Model 1 (One Spike)</th>
<th>Model 2 (Two Spikes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Two Parameters -- m, n</td>
<td>f_1(p) is beta distributed with parameters -- m, n</td>
<td>Three Parameters -- α, m, n</td>
<td>Four Parameters -- α, β, m, n</td>
</tr>
<tr>
<td></td>
<td><img src="image4" alt="Equation 1" /></td>
<td><img src="image5" alt="Equation 2" /></td>
<td><img src="image6" alt="Equation 3" /></td>
</tr>
<tr>
<td>p_0</td>
<td>(p_j)_{j=0}</td>
<td>α + (1-α) \frac{B(m+j, n+k)}{B(m,n)} + (1-α-β) \frac{B(m+n+k)}{B(m,n)}</td>
<td>(1-α-β) \frac{B(m+j, n+k-j)}{B(m,n)}</td>
</tr>
<tr>
<td>p_j</td>
<td>c_j \frac{B(m+j, n+k-j)}{B(m,n)}</td>
<td>(1-α)c_j \frac{B(m+j, n+k-j)}{B(m,n)}</td>
<td>(1-α-β) \frac{B(m+k,n)}{B(m,n)}</td>
</tr>
<tr>
<td>p_k</td>
<td>(p_j)_{j=k}</td>
<td>(p_j)_{j=k}</td>
<td>(1-α-β) \frac{B(m+n+k)}{B(m,n)}</td>
</tr>
<tr>
<td>Average Purchase Frequency</td>
<td>\bar{j} = (\frac{m}{m+n}) k</td>
<td>\bar{j} = (1-α) (\frac{m}{m+n}) k</td>
<td>\bar{j} = (1-α-β) (\frac{m}{m+n} + β) k</td>
</tr>
</tbody>
</table>
a monotonic transform of $\ell(M)$ would not change the parameter estimates for which the maximum is attained, we replace the constant term in $\ell(M)$ by unity and work with the logarithm of the altered likelihood function for computational convenience. Thus, the maximum likelihood estimate of $\theta$ is obtained by maximizing $L(M)$, which is a monotonic transform of $\ell(M)$, where:

$$L(M) = \log\{(p_0)^{n_0} (p_1)^{n_1} \ldots (p_k)^{n_k}\},$$

or,

$$L(M) = \sum_{j=0}^{k} n_j \log p_j.$$  

The probabilities $p_j (j = 0,1,\ldots k)$ in the expression are functions of the parameters of the model as shown in Table 1. For instance, in the case of Model 2, the $p_j$'s would depend on the parameters of this model — $m, n, \alpha$ and $\beta$; the maximum likelihood estimates of the parameters are obtained by maximizing $L(2)$ with respect to $m, n, \alpha$ and $\beta$. The expression for $L(M)$ in equation (2) above on substitution for $p_j$ from Table 1, becomes quite complex and, hence, an analytical derivation of the maximum likelihood estimates of the parameters is not feasible. Therefore, numerical optimization is used to obtain the values of the parameters. The computer program developed for this purpose — a modified pattern search — is based on the pattern search process developed by Hookes and Jeeves [3]. The program provides an optimization procedure for determining the values of $n$ different parameters at which a given function of these $n$ parameters attains a global optimum value. This modified pattern search program was developed by Kalwani for his doctoral dissertation [4].
THE LIKELIHOOD RATIO TEST

The likelihood ratio test uses the maximum value of the likelihood function for each of the two models. One model must be a constrained version of the other model. Model 2 which contains two spikes in its purchase probability distribution is unconstrained and has four parameters (α, β, m and n); two of these four parameters are the parameters of the beta distribution (i.e., m and n) and the other two parameters represent the two spikes, α and β. Model 1 which has one spike is a constrained form of Model 2; it has three parameters (α, m and n). For Model 1 the spike parameter β, representing the proportion of "always buy" consumers is constrained to take on the value zero. In the case of Model 0, both the spike parameters, α and β, are restricted to take on the value zero. Clearly,

\[ \text{Model 0} \subset \text{Model 1} \subset \text{Model 2}. \]

Therefore, \( L^*(0) \leq L^*(1) \leq L^*(2) \).

Where \( L^*(M) \) represents the maximum value of the monotonic transform of the likelihood function (see equation (2)) for Model 0, 1 or 2. While it is obvious that \( L^*(0) \) is less than \( L^*(1) \) or \( L^*(2) \), the question is whether or not this difference is statistically significant. That is, if we set up the constrained model as the null hypothesis \( (H_0) \) and the unconstrained model as the alternate hypothesis \( (H_1) \), we would like to see if the data suggest that we reject the null hypothesis in favor of the alternate hypothesis. The likelihood ratio test answers this question. We have:

\[ H_0: \text{ Model 0 (or 1) with } \omega=2 \text{ (or 3) parameters} \]
\[ H_1: \text{ Model 1 (or 2) with } \Omega=3 \text{ (or 4) parameters} . \]
Next, we form a ration (say, $\lambda$) of the maximum value of the likelihood function of the constrained model under the null hypothesis (say, Model 0) to that for the unconstrained model (say, Model 2).

$$\lambda_{02} = \frac{\ell(0)}{\ell(2)}.$$  

For large samples, Wilks [11] has shown that, $-2 \log_e \lambda$ is distributed chi-square with $(\Omega - \omega)$ degrees of freedom. In testing Model 2 against Model 0, since $L^*(M)$ from equation (2) represents the logarithm of the maximum value of the likelihood function for Model M, we have:

$$-2 \log_e \lambda_{02} = -2[L^*(0) - L^*(2)] = 2[L^*(2) - L^*(0)] \quad (3)$$

is distributed chi-square with 2 (=4-2) degrees of freedom.

**PURCHASE SEQUENCE LENGTH AND SAMPLE SIZE**

In this section we determine the sample size and the purchase sequence length required to detect spikes. Obviously a sample of 50 consumers with three purchases each is not sufficient to test if a proportion of the consumers in the population are "always buy" or "never buy" customers. On the other hand, a sample of 10,000 consumers with 20 purchases is clearly adequate to test for spikes. It is not obvious what purchase sequence lengths and sample sizes are required to detect mass-points in purchase probability distributions. Also, it would be useful to know the trade-off between sample size and purchase sequence length requirements for detecting and estimating spikes.
We determine these requirements through the use of simulated data. We generate samples from eleven different settings which extend over the spectrum of zero-order models. The first five settings are based on purchase probability distributions with no spikes. The probability distribution is allowed to take various shapes like uniform, bell, "decay curve," U and reverse J. These five settings are described in greater detail with the actual values of the parameters in a different paper by the authors (see Kalwani and Morrison [5]).

The next three settings from which simulated sample data are generated contain skewed bell shape distributions with spikes at \( p=0 \). The spikes representing "never buy" consumers are of size 5%, 10% and 20% respectively. The last three settings are based on bell shape purchase probability distributions with two spikes at \( p=0 \) and \( p=1 \). The two spikes are of equal size in each of the three cases and are of size 5%, 10% and 20%.

Many researchers have used sample data with purchase sequence lengths of 5 or less for estimating parameters of their models. There has been, however, considerable variation in the sample sizes used for parameter estimation. To start with, we begin by generating samples of size 500 with a purchase sequence length of 5 for each of the eleven settings (5 Model 0, 3 Model 1 and 3 Model 2). These sample specifications, as we shall see in the next section, prove to be adequate for the first five settings of models with no spikes. For the one and two spike models, however, these sample specifications turn out to be inadequate and we repeat the analysis by increasing the sample size to 1000 with a purchase sequence length of 5 and then by increasing the purchase sequence length to 10 for samples of size 500.
The testing procedure for the simulated data consists of generating 10 samples at each preset sample specification (say, sample size = 500 and purchase sequence length = 5). The likelihood ratio test is performed on the 10 simulations. The likelihood ratio test determines the model (no-spike, one-spike or two-spike) which forms the best fit for the simulated sample data obtained from each of the 10 simulation runs. From this testing of the simulated data we determine the number of times (out of 10) the true model (no-spike, one-spike or two-spike) that is used to generate the sample data is correctly identified as providing the best fit for the data. Sample specifications in terms of sample size and purchase sequence length are adequate if the true model can be correctly identified most of the time with these specifications.

Let us now analyze the findings from this testing of the simulated data which is performed at a confidence level of 95%. Sample specifications of sample size = 500 and purchase sequence length = 5 are found to be adequate for no-spike models. That is, for the sample data generated from the first five settings (no-spike models) the null hypothesis of Model 0 with no spikes is never rejected in favor of the alternate hypothesis of Model 1 with one spike or Model 2 with two spikes.

Table 2 displays the results obtained in the case of one-spike and two-spike models. As mentioned earlier the likelihood ratio test is performed for three different sample specifications in case of the spiked models. Under each of the three sample specifications Table 2 displays for each setting the number of times out of 10 that a particular model provides the best fit for the simulated data. The number of times for which the true model is correctly identified is underlined. As an illustration of the results
<table>
<thead>
<tr>
<th>Setting #</th>
<th>Sample Size = 500</th>
<th>Sample Size = 1000</th>
<th>Sample Size = 500</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mod. 0</td>
<td>Mod. 1</td>
<td>Mod. 2</td>
<td>Mod. 0</td>
</tr>
<tr>
<td>6. one-spike $\alpha = 0.05$</td>
<td>9</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>7. one-spike $\alpha = 0.10$</td>
<td>7</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>8. one-spike $\alpha = 0.20$</td>
<td>4</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>9. two-spike $\alpha = 0.05$</td>
<td>8</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$\beta = 0.05$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 two-spike $\alpha = 0.10$</td>
<td>3</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>$\beta = 0.10$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11 two-spike $\alpha = 0.20$</td>
<td>3</td>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td>$\beta = 0.20$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

displayed in Table 2, the findings for the setting 7 show that with the sample specifications of size 500 and purchase sequence length of 5, the true model (i.e., Model 1 with one spike) is correctly identified only three times out of ten. Doubling the sample size to 1000 with the same purchase sequence length raises the number of correctly identified models to 7 whereas doubling the purchase sequence length leads to correct identification of the true model all the ten times. The remainder of Table 2 can be interpreted in the same manner for settings 8 through 11.
Several interesting and useful conclusions can be drawn from the findings in Table 2. The first rather obvious inference we make is that it is easier to detect larger spikes both for one-spike and two-spike models.

For the sample specifications of size 500 and purchase sequence length of 5, the number of times the true model is correctly identified are 9 (out of 30) for the one-spike models and 13 for the two-spike models. Raising the sample size to 1000 (the purchase sequence length is still 5) increases these numbers to 18 and 23 respectively. The gain in spike detecting power is much greater if the purchase sequence length rather than the sample size is doubled. For sample size = 500 and purchase sequence length = 10, the number of times the true model is detected are 28 for the one-spike models and 30 for the two-spike models. We conclude that doubling the purchase sequence length from 5 to 10 gives more additional power to detect spikes rather than doubling the sample size from 500 to 1000.

Overall a purchase sequence length of 10 with sample size of 500 or larger is adequate to identify the true model (58 out of 60 times for models with one or two spikes). We conclude that a purchase sequence length of 10, which for many products and most consumers represents purchases of branded goods over half a year, represents a desirable purchase sequence length to detect and estimate the proportion of "always buy" and/or "never buy" consumers.
EMPIRICAL DATA

Description:

The empirical data used in this paper come from two sources. The first data set is from Aaker [1] who obtained it from the MRCA panel. This sample only includes new triers of a particular brand of a frequently bought good. The second source for the empirical data is from Montgomery [8] and once again the original source is the MRCA panel. The data are from the dentrifice market just prior to and immediately following the endorsement of Crest toothpaste by the American Dental Association. Both the before and after endorsement data are used in this paper. Montgomery's model and a description of the data can be found in Montgomery [9]. However, the specific numbers in Table 3 are only in the unpublished working paper, Montgomery [8].

All the three data sets are for a purchase sequence length of 5 and the purchase frequencies are displayed in Table 3. In view of the findings from the simulated data in the previous section we should note that the sample size of 631 (with 5 purchases per consumer) for Aaker's data is not sufficient to test for spikes.

TABLE 3

<table>
<thead>
<tr>
<th># of purchases of Brand 1</th>
<th>Aaker's Data</th>
<th>Montgomery's &quot;Before Endorsement&quot; Data</th>
<th>Montgomery's &quot;After Endorsement&quot; Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>.352</td>
<td>2,212</td>
<td>1,683</td>
</tr>
<tr>
<td>1</td>
<td>125</td>
<td>383</td>
<td>436</td>
</tr>
<tr>
<td>2</td>
<td>62</td>
<td>154</td>
<td>270</td>
</tr>
<tr>
<td>3</td>
<td>38</td>
<td>91</td>
<td>203</td>
</tr>
<tr>
<td>4</td>
<td>30</td>
<td>89</td>
<td>209</td>
</tr>
<tr>
<td>5</td>
<td>24</td>
<td>106</td>
<td>234</td>
</tr>
<tr>
<td>TOTAL</td>
<td>631</td>
<td>3,035</td>
<td>3,035</td>
</tr>
</tbody>
</table>
Findings:

Our objective is to find which one of the three models (namely Models 0, 1 or 2), on the basis of the likelihood ratio test, provides the best fit. In order to perform the likelihood ratio test we need to calculate the logarithm of the maximum value of the likelihood function for each of the three models. Computer optimization through the modified pattern search program provides the logarithm of the likelihood function for each of the models and the maximum likelihood estimates of the parameters. The results from these computer runs are displayed in Table 4.

Table 4 displays in rows I the maximum likelihood estimates of the Model parameters -- m, n, α and β. The next set of rows II contain the logarithm of the maximum value of the likelihood function. The last set of rows III display the test statistic, $-2 \log_{e} \lambda$, for each of the data sets. For instance, $-2 \log_{e} \lambda_{02}$ is the test statistic when Model 0 is the null hypothesis and Model 2 is the alternate hypothesis; equation (3) is used to compute this test statistic.

Examination of the test statistic values reveals that for Aaker's data null hypothesis of Model 0 is not rejected in favor of either Model 1 or 2; we conclude that Model 0 provides the best fit for Aaker's data (but recall the inadequacy of his sample size). In the case of Montgomery's "Before Endorsement" data Models 0 and 1 are rejected in favor of Model 2; hence Model 2 forms the best fit to the second data set. For the third data set Model 0 provides the best fit since it is not rejected in favor of Model 1 or 2. In fact it is a little disconcerting that for this third set of data the addition of spikes has absolutely no effect on the likelihood functions.

The most significant finding in Table 4 is that Model 1 does not constitute a best fit in any of the three cases (see test statistic values, $-2 \log_{e} \lambda_{01}$). This result in view of the proportion of the consumers with no purchases of Brand 1 is nonintuitive and warrants an explanation. For
\[
\begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|c|}
\hline
\text{Data Set} & \text{Aaker's Data} & \text{Montgomery's "Before Endorsement" Data} & \text{Montgomery's "After Endorsement" Data} \\
\text{Parameters} & m & n & \alpha & \beta & m & n & \alpha & \beta & m & n & \alpha & \beta \\
\hline
\text{I} & & & & & & & & & & & & \\
\text{Model 0} & 0.336 & 1.401 & 0 & 0 & 0.142 & 0.985 & 0 & 0 & 0.234 & 0.751 & 0 & 0 \\
\text{1} & 0.336 & 1.397 & 0.000 & 0 & 0.142 & 0.985 & 0.000 & 0 & 0.234 & 0.751 & 0.000 & 0 \\
\text{2} & 0.382 & 1.772 & 0.000 & 0.018 & 0.173 & 1.512 & 0.000 & 0.022 & 0.234 & 0.752 & 0.000 & 0.000 \\
\hline
\text{II} & & & & & & & & & & & & \\
\text{L*(0)} & -830.65 & & & & -2959.47 & & & & -4200.96 & & & \\
\text{L*(1)} & -830.65 & & & & -2958.73 & & & & -4200.96 & & & \\
\text{L*(2)} & -829.37 & & & & -2948.48 & & & & -4200.96 & & & \\
\hline
\text{III} & & & & & & & & & & & & \\
\text{\(-2 \log \lambda_{01}\)} & 0.00 & & & & 1.48 & & & & 0.00 & & & \\
\text{\(-2 \log \lambda_{02}\)} & 2.56 & & & & 21.98 & & & & 0.00 & & & \\
\text{\(-2 \log \lambda_{12}\)} & 2.56 & & & & 20.50 & & & & 0.00 & & & \\
\hline
\end{array}
\]

Note: The maximum likelihood estimates of the parameters in rows I have been rounded to the third decimal place and the logarithms for the values of the maximum likelihood function in rows II have been rounded to the second decimal place.
the three data sets displayed in Table 3 the proportions representing the consumers who make 0 purchases of Brand 1 are 56%, 72% and 55% respectively. An examination of the proportions at the other purchase frequencies shows that the shape of the purchase probability distribution representing these data sets is likely to be either U-shaped or concave (when \( m \) and \( n \) are both less than 1) or a shape where the curve steadily declines (when \( m < 1 < n \)). This observation is confirmed by the maximum likelihood estimates of the parameters of the beta distribution (\( n \) and \( n \)) displayed in Table 4 for Model 0.

These two shapes are displayed in Figures 1 and 2. The proportion of consumers who make one or more purchases of Brand 1 constrain the purchase probability distribution to take one of the two shapes shown in the figures above. To detect a spike on these purchase probability distributions is going to be difficult as is illustrated in Figure 3. In fact, many researchers would argue that Figure 3 is essentially equivalent to Figure 1. On the other hand, a spike at \( p = 0 \) would be easier to detect for the case shown in Figure 4.
We conclude that it is the frequencies for non-zero purchases which determine the shape of the purchase probability distributions. Our findings have support in the work of marketing practitioners. According to Butler [2] and Kropp [6] the purchase probability distributions for mature brands are generally described by U-shaped distributions. Stewart [10] in an empirical study of distributions of purchase proportions states that, "the distribution of proportions is U-shaped for virtually all brands in each field irrespective of their size." Note that the Aaker data which is for a new product gives a steadily decreasing J (or non-U) shaped distributed. Montgomery's data on the established Crest brand gives two U-shaped distributions.

CONCLUSION

In this paper, we have presented, the use of the likelihood ratio test as a natural method to estimate for the presence of "always buy" and "never buy" consumers.

Purchase sequence length and sample size requirements were determined through the use of simulated data. We found that doubling the purchase sequence length from 5 to 10 gives more additional power to detect spikes rather than doubling the sample size from 500 to 1000. In other words for spike detecting power it is useful to trade-off an increase in sample size for an increase in purchase sequence length. As a rule of thumb, a purchase sequence length of 10 is needed to detect spikes especially if the sample size is less than 1000. Not many published studies in the marketing literature have had this large a purchase sequence length. Most researchers have used purchase lengths of 5 or less.
Our findings show that although the empirical data contained more than 50% consumers who had never bought a given brand the specific addition of a spike to provide for "never buy" consumers did not form a superior fit over the no-spike model. The three data sets used here provide additional support for U-shaped purchase probability distribution found by some marketing practitioners. However, it is not fair to conclude that these three data sets included no "never buy" or "always buy" consumers. As we have seen the purchase sequence length of five used in these studies is inadequate to detect spikes. We merely have shown that a beta distribution without spikes can fit these data sets as well as the more general distribution that includes spikes.
REFERENCES


