The Impact of Promotion and Advertising on
Choice of Fruit Category and Apple Variety: A Latent Class Approach

by

Timothy J. Richards and Paul M. Patterson¹

Paper presented at AAEA Annual Meetings

Nashville, TN. August, 1999

¹ Associate and Assistant Professors, Morrison School of Agribusiness, Arizona State University East. The authors wish to express their appreciation for helpful comments by Henry Kinnucan, John Nichols, Ram Acharya and the participants in the NEC 63 Commodity Promotion Research Conference, Oct. 5-6, 1998 in Tempe, AZ. Funding from the National Institute for Commodity Promotion and Research (NICPRE) is also gratefully acknowledged. Copyright 1998 by Timothy J. Richards and Paul M. Patterson. All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.
Abstract

With the increasing sophistication of consumer tracking information technology, marketing managers are becoming better able to segment consumers, thereby increasing the efficiency of their promotional programs. Typically, marketers define segments using demographic characteristics, implicitly believing that these segments represent consumers with relatively homogeneous buying patterns. A more managerially useful definition of a market segment, however, groups consumers of similar behavior directly and then seeks to find demographic commonalities among them. This study uses a latent class analysis technique to segment consumers based on their responsiveness to a set of marketing variables (price, advertising, and promotion). To help target the more responsive segments, ex post demographic analysis finds relationships between advertising responsiveness and a set of demographic characteristics. Using A.C. Nielsen panel scanner data for fruits, this study finds that a multiple-segment model provides a better fit to the data, and that these segments differ significantly in their responsiveness to marketing variables. By targeting marketing activities to their most responsive segments, the effectiveness and efficiency of commodity promotion can be dramatically improved.

Keywords: advertising, apple, demand, latent class analysis, multinomial logit, segmentation
The Impact of Promotion and Advertising on
Choice of Fruit Category and Apple Variety: A Latent Class Approach

Introduction

By segmenting a market into groups of relatively homogeneous consumers, marketing managers hope to be able to increase the efficiency of promotional programs directed towards those segments. However, the usefulness of this method relies upon the implicit assumption that consumers that share similar demographic profiles also share similar buying habits. In many cases, this is clearly a very strong assumption and one that is not likely to hold up to deeper scrutiny. It would seem more logical, if behavior is of primary interest, to segment consumers on the basis of similar behaviors and then search for any demographic commonalities among them. The ability to segment consumers directly into groups that respond similarly to various marketing tools (price, promotion, advertising) is even more powerful if these segments are defined specific to their decision-stage (whether or not to buy from a product category on a particular shopping trip and what brand within the chosen category).

Because unique response-segments are likely to exist at each stage, it is necessary to be able to both estimate the size and membership in these segments as well as estimate the response parameters in each (Bucklin, Gupta, and Siddarth). Several studies demonstrate the usefulness in estimating market segment sizes and unique segment-specific response parameters using household scanner data for consumer packaged-product purchases (Grover and Srinivasan (1987); Kamakura and Russell; Bucklin and Lattin; Bucklin and Gupta; Grover and Srinivasan (1992); Krishnamurthi and Raj; Bucklin, Gupta and Han; Bucklin, Gupta and Siddarth). However, none of this work addresses the unique characteristics of fresh fruit. Perishability, frequency of purchase, variable quality and a lack of true brands are each likely to influence consumers’ choice between varieties of a particular fruit and the decision to buy one fruit versus another. This type of analysis is now possible with the development of household scanner panel data sets for what were once non-scanable products.
The objectives of this research are, therefore, to first determine whether there exists segments of fruit consumers whose members respond similarly to marketing variables such as price, promotion, and advertising. If such segments exist, a second objective is to determine the factors that drive the choice of a particular type of fruit (category choice) and fruit variety (variety choice) within each segment. By estimating the probability that each household belongs to each segment, this research also determines segment membership and, using household demographic data, describes each segment’s typical demographic profile.

**An Empirical Model of Variety Choice, Category Choice and Market Segmentation**

The underlying rationale for separating variety choice from category choice, or the timing of category purchase, is the realization that consumer behavior is likely to differ according to the stage of the decision-making process. In particular, many argue (Tellis; Bucklin and Gupta; Dillon and Gupta) that the probability of a consumer buying within a category of products (i.e., bananas or apples) depends upon a different set of factors than those that determine the probability of choosing a particular variety within a category (i.e., Red Delicious apples). While the former may be determined by a real or perceived need as measured by such things as family size, inventory level, or the intensity of category advertising, the latter is more likely to reflect prices, promotion, tastes or loyalties. Therefore, this study adopts a nested multinomial logit approach in estimating elasticities of variety and category choice. Tastes and loyalties reflect not only differences in behavior between decision stage, but also between heterogenous households. To account for this possibility, the model estimates separate elasticities for multiple segments of consumers.

Being able to identify these segments promises not only a more accurate characterization of the market, but is also far more useful to marketing managers who are attempting to target consumer segments for various marketing campaigns. The problem is, therefore, how to identify these segments? Because consumers’ response parameters are unobservable *a priori*, segments of homogeneous
consumers are also unobservable, or “latent classes” (Grover and Srinivasan 1987; Lazarsfeld and Henry). Treating the probability that a household belongs to each latent class as a parameter to be estimated, this study uses a finite mixture estimation (FME) method (Titterington, Smith, and Makov) to find these probabilities, and hence, segment membership. In a FME, or mixture model, rather than assuming the errors are drawn from one distribution, they are instead allowed to belong to two or more distributions, each distribution corresponding to a segment. The likelihood of a particular observation is, therefore, a weighted average of the likelihoods finding the observation in each of the segments. By maximizing the sum of these likelihoods, the technique yields both segment-specific response parameters and estimates of each segment share, or segment probabilities.

Within each segment, a household’s choice of apple variety is modeled using McFadden’s (1974) random utility assumption. While this study follows Grover and Srinivasan (1987, 1989, 1992); Bucklin and Gupta; and Dillon and Gupta in developing a model of choice where consumers belong to one of several mutually exclusive response segments, the economic model initially assumes only one segment for clarity of exposition. For this single response-segment, the random utility model assumes that a consumer faces a choice between goods that are effectively perfect substitutes. With linear indifference curves, the usual income-constrained utility maximization process guarantees that only one of the goods will be purchased (Deaton and Muellbauer; Pudney). The empirical model, therefore, seeks to explain the probability that a consumer chooses the particular variety \( i \in I \) of fruit \( k \in K \) that provides maximum utility from discrete alternatives within each set of \( I \) (varieties) and \( K \) (fruits). If consumers choose sequentially from among product categories, and then from within each category according to the multi-stage budgeting process of Gorman or McFadden, the total probability that a consumer chooses variety \( i \) from category \( k \) is the product of the conditional probability of choosing \( i \) given \( k \) and the marginal probability of choosing category \( k \). Therefore, the unconditional probability of household \( h \) choosing variety \( i \) from category \( k \) at time \( t \) is given by:
\[ P_t^h(i, k) = P_t^h(i|k) P_t^h(k), \]  

(1)

where \( P_t^h(k) \) is the probability of buying fruit from category \( k \) on a given shopping trip \( t \), and \( P_t^h(i|k) \) is the probability of purchasing variety \( i \) conditional on the choice of category \( k \). Because the probability of each variety and category choice is determined by the relative utility derived from each compared to its alternatives, implementing (1) requires the definition of formal random utility models at each decision stage. Assuming that the household’s utility is composed of a deterministic and random component, the total utility from both choices is given by:

\[ U_t^h(i, k) = \{ V_t^h(i, k) + \epsilon_t^h \} = u_t^h + X_t^h + Y_t^h + \epsilon_t^h, \]  

(2)

where \( u_t \) represents a choice-specific preference parameter, \( X_t^h \) consists of a set of household and variety attributes that can vary by product category, \( Y_t^h \) is a set of category-specific variables, and \( \epsilon_t^h \) is a random error term. Assuming this error term is type I extreme value distributed at each stage (McFadden 1981), the conditional variety-choice probability in (1) is written as:

\[ P(i|k) = \frac{\exp(i X_{ik})}{\exp(I_k)}, \]  

(3)

while the probability of choosing category \( k \) is given by:

\[ P(k) = \frac{\exp(k Y_k + I_k)}{\sum_{j=1}^{T} \exp(j Y_j + I_j)}, \]  

(4)

where \( I_k = \log(\sum_{m=1}^{N} \exp(m X_{km})) \) is defined as the “inclusive value” or, in this application, the “category value” term (Ben-Akiva and Lerman; Bucklin and Gupta) for \( T \) categories and \( N \) varieties.
Including this term in the category choice model insures that the product of the variety and category choice probability models represents the unconditional probability of variety purchase. Such disaggregation provided by the nested multinomial logit model offers other attractive features beyond theoretical consistency.

Whereas the choice model in (1) assumes that all households share a common vector of response parameters, assume instead that there are a finite number of consumer segments that are relatively homogeneous in terms of both their variety \((r)\) and category \((s)\) response. In this case, the unconditional probability that a household buys variety \(i\) from category \(k\) in (1) can be rewritten as a weighted average, or mixture of the variety and category choices across segments of each (Bucklin and Gupta):

\[
P^h(i, k) = \sum_r \sum_s r_s P^h_r(i | k) P^h_s(k),
\]

where \(r_s = \text{the size of the } rs \text{ variety choice and category choice segment}, 0 \leq r_s \leq 1\), and \(\sum_r \sum_s r_s = 1\). Because these segment probabilities, or latent classes, are unobservable, the EM (expectation / maximization) algorithm described in Dempster, Laird, and Rubin is used to estimate them using household-panel scanner data.

**Variable Definition and Data Sources**

At the variety-choice level households choose from among a set of apple varieties, conditional on their choice of the apple category. These varieties include Red Delicious, Golden Delicious, Granny Smith, and a set of other or “specialty varieties.” In making this choice, the deterministic component of a household’s utility from the first-stage or variety choice, \(V^h_i\), consists of variables reflecting attributes of both the household and its particular choice:
\[ V_i^h = u_i + \beta X_i^h \]
\[ = u_i + \beta_{PR} I_{PR}^h + \beta_{DL} I_{DL}^h + \beta_{LP} I_{LP}^h + \beta_{LOY} I_{LOY}^h + \beta_{LOYPR} I_{LOYPR}^h + \beta_{LOYDL} I_{LOYDL}^h, \]  

where:  
- \( PR \) = the adjusted price of choice \( i \) facing the household;  
- \( DL \) = binary variable indicating whether the product is on a promotional deal;  
- \( LP \) = binary variable indicating whether the last purchase was the same as the current purchase;  
- \( LOY \) = binary variable indicating whether the household is variety loyal;  
- \( PR \cdot LOY \) = interaction variable between loyalty and price;  
- \( DL \cdot LOY \) = interaction variable between loyalty and deal.

This specification is similar to Bucklin and Gupta, but reflects many features unique to the fruit-choice data. First, prices for similar varieties are commonly assumed to be identical for sample households. Although this may be a valid assumption for single-market household samples buying national brands of packaged consumer products, this study includes households from six major U.S. metropolitan markets buying fruit that often vary in local quality, availability, size, grade and outlet. To control for the variability in price due to unobservable or “quality” characteristics, this study uses the hedonic method of Cox and Wohlgenant to create quality adjusted fruit prices. Further, all prices are also net of any promotional deals so the reported price is that which is actually paid. This adjustment also allows for the inclusion of a separate variable designed to capture the independent effect of price-promotions on variety choice.

Second, models of variety and category choice typically find that loyalty is the most important explanatory variable (Guadagni and Little; Tellis; Gupta; Bucklin and Lattin). Operationalizing the notion of “loyalty,” however, is subject to considerable controversy. In this study, we adopt the approach of Krishnamurthi and Raj (1991) who define a household as loyal if its dominant brand share is greater than 50%. Because this definition of loyalty is constant over the entire purchase-history of
each household, it is often thought of as a long-term measure of loyalty. Grover and Srinivasan (1992) and Bucklin and Gupta make compelling arguments for including a second, short-term measure of loyalty. Thus, LP is a binary indicator of variety loyalty that is assigned a value of 1.0 if a consumer buys the same variety on two adjacent trips to the store. Research that explicitly defines loyal and non-loyal buyers commonly finds a significant difference in price response between the two groups, so we include interaction terms between these variables as well. In contrast to these preference variables, category choice tends to follow patterns of need.

To capture these effects, the vector $Y$ in the category model includes few variables in common with the vector of variety-choice determinants. Specifically, the category-utility for household $h$ is written as:

$$ V_{k}^{h} = Y_{k}^{h} = \begin{align*} &1 \cdot LP_{k} + 2 \cdot CR_{k} + 3 \cdot INV_{k} + 4 \cdot ALOY_{k} + 5 \cdot AAD_{k} + 6 \cdot OAD_{k}, \end{align*} \tag{7} $$

where the explanatory variables introduced here include:

$CR =$ consumption rate, in pounds per day ;

$INV =$ household inventory ;

$ALOY =$ category loyalty ;

$AAD =$ apple advertising ;

$OAD =$ other fruit advertising.

Including the last purchase and loyalty variables accounts for both the short- and long-term notions of loyalty defined above for the variety model. As in the variety model, we expect both of these variables to be significant, positive determinants of category choice. For purposes of the category model, a household is “category loyal” if it purchases more than 50% of its fruit needs from either the apple, banana, grape, or soft fruit categories. Further, this study follows Bucklin and Gupta by calculating inventory from:
\[
INV^h_t = INV^h_{t-1} + Q^h_{t-1} + CR^h I_{t,t-1},
\]

where \(INV^h_t\) is the inventory of household \(h\) at time \(t\), \(Q^h_{t-1}\) is the amount purchased on the previous shopping trip, \(CR^h\) is the average daily category consumption rate, calculated from the entire purchase history of household \(h\), and \(I_{t,t-1}\) is the interval between successive purchases, measured in days.

Whereas each of these variables represent an actual need for the product by each household, advertising reflects a household’s perceived need for the product category.

Although key to category choice, variety choice decisions are unlikely to be influenced by generic advertising. For purposes of this study, advertising includes the total amount of market-specific expenditure from all sources on television, radio, newspapers, and outdoor advertising as reported by Competitive Media Reporting, Inc. (CMR). Because CMR reports their findings on a monthly basis only, and the unit of observation here is the shopping-trip day, the amount of advertising exposure is assumed to be constant for all shopping trips taken within a particular month.

Except for the CMR data, all data are from the AC Nielsen HomeScan database collected over the period of July - December, 1997. For this study, the category sample consists of observations covering over 38,000 shopping trips wherein consumers bought either apples, bananas, grapes, or soft fruit (nectarines, peaches, plums, or pears). Although the data include many other types of fruit, they are not in the final sample due to infrequency of purchase. This sub-sample consists of 9,510 purchase occasions after excluding households that make less than two fruit purchases per month. Although this may be a source of sample-selection bias, it is common in this literature and necessary to implement the loyalty and inventory variables with any degree of confidence. Each observation includes the date of the purchase, price paid, quantity purchased, use of a promotional deal, shopping outlet, metropolitan area,
and a variety of household demographic data. With this data, all estimates are obtained using maximum likelihood routines in Shazam 7.0.

**Results and Discussion**

At both the variety and category levels, likelihood ratio (LR) tests compare the fit of a two segment to the single-segment alternative. In the variety model, a LR of 13,249.38 strongly supports the superiority of the two-segment model. The results show that segment one consists of over 62% of the sample households with the remainder in segment two, so it may indeed be both feasible and profitable to attempt to reach both.

In a multinomial logit model, the structural parameters do not have intuitive interpretations, therefore, table 1 shows the elasticities of variety choice. These elasticities suggest that segment one is both more price elastic and deal-sensitive than segment two. Although households in segment one tend to switch varieties in response to differences in price, they are also more likely to exhibit strong short- and long-term variety loyalty. Moreover, this result implies that segment one households may be easier to attract with low prices, but once they buy a particular variety habitually, they are less likely to switch than segment two members. These differences are particularly valuable if it is possible to identify the segments by various demographic characteristics, and if these characteristics are also common to category choice segments. [table 1 in here]

As the chi-square likelihood ratio test statistic in table 2 shows, a two-segment model represents a significant improvement over the single-segment alternative. Again, the category choice elasticities show that consumers are far more short-term loyal than long-term loyal. Category managers,

---

1 The likelihood ratio test statistic is: \( LR = 2(LLF_u - LLF_r) \), where \( LLF_u \) is the value of the unrestricted log-likelihood function (two-segments) and \( LLF_r \) is the value of the restricted log-likelihood function. With 28 parameter restrictions, the critical chi-square value at a 5% level is 41.337.
therefore, should focus more resources on initiating new purchases, rather than attempting to reinforce established patterns. Segment one purchases are also more sensitive to the rate of household consumption, but less so with respect to inventories. Consistent with their pattern as the “sensitive segment,” households in segment one are also more responsive to mass apple-advertising, but both segments are equally influenced by non-apple advertising. However, this difference is only useful if a manager is able to identify segment members and direct marketing efforts appropriately. [table 2 in here]

As a final step in the analysis, we characterize the demographics each response-segment using household-specific data describing household size, income, age, education, and metropolitan market. With respect to variety choice, households in the “more responsive” segment (one) have significantly higher incomes, smaller households, are older, and are slightly less educated than their less-responsive counterparts. Combining this result with the response elasticities invites characterizing a typical household in the first segment as an “easy sell,” one that is highly receptive to incentives to change varieties, and more prone to continue buying a variety of apple that it likes. Geographically, consumers of the first type are disproportionately located in New York and San Antonio, as opposed to Chicago and Los Angeles for members of the second segment. These profiles differ slightly from those defined for the category choice model. Whereas households that are more responsive to marketing and need variables in category choice are again smaller, older, and less educated than those in the less-responsive segment, in this case they have significantly lower incomes. Given the difference in average age and household size of members in these two segments, households that may be relatively easy to switch from some other type of fruit to apples may not have as much discretionary income as those that are willing to experiment with new varieties. The importance of the short-term definition of loyalty to these people suggests that free samples or demos would be effective in building category volume. Further,
and perhaps more importantly, apple advertising, which appears to be able to overcome the competing effects of banana advertising, can be effective in building the apple category if it induces initial purchases.

**Conclusion and Implications**

The central hypothesis of this paper is that marketing activities, particularly price-promotions and media advertising, have different effects on consumers’ choice of fruit-variety compared to their effect on consumers’ decision to buy a particular type or category of fruit on a given shopping trip. This study tests this hypothesis using a nested multinomial logit (NMNL) model that estimates elasticities of apple variety and category choice specific to heterogeneous segments of consumers. This latent class model is estimated using a maximum likelihood EM approach with over 9500 fruit-purchase observations gathered from a large sample of A.C. Nielsen scanner-enabled households on a shopping-trip basis over the final six months of 1997.

With this approach, the paper finds a significant value to differentiating between factors that drive consumers’ variety and category choices. Whereas factors such as variety-specific preferences and prices are important in a household’s choice of apple-variety, need-based factors such as consumption rate and level of household inventory are more important in determining whether consumers purchase apples as opposed to another type of fruit on a particular trip to the store. In addition to the measures of actual need, as an indicator of perceived need, mass-advertising is a significant determinant of the probability that a household buys apples on a given shopping trip.

The results also show that a two-segment model provides a far better fit to the data than a single-segment version at each decision-stage. At the variety-choice stage, households in one segment are shown to be more responsive to changes in price and individual measures of variety-loyalty
compared to the other. A typical household in this segment is smaller, has a higher income, is older, and
has less education than one in the less responsive segment. At the category level, members of the more
responsive segment share the household size, age, and education characteristics of the variety-choice
responsive segment, but have lower incomes than the less-responsive segment. Membership in these
segments also differs significantly by market. These segment-specific elasticities will allow commodity
marketers to target their limited promotion budgets more efficiently and to make better use of the
household-specific purchase data that is now.
Reference List


Gupta, S. “Impact of Sales Promotions on When, What, and How Much to Buy.” *Journal of Marketing*


Table 1. Variety Choice Elasticities by Response Segment

<table>
<thead>
<tr>
<th>Var:</th>
<th>Choice Segment 1</th>
<th>Choice Segment 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Red,$^1$</td>
<td>Gold</td>
</tr>
<tr>
<td>PR</td>
<td>-1.129</td>
<td>-0.338</td>
</tr>
<tr>
<td>LP</td>
<td>0.392</td>
<td>0.241</td>
</tr>
<tr>
<td>LOY</td>
<td>2.172</td>
<td>1.775</td>
</tr>
<tr>
<td>DL</td>
<td>0.112</td>
<td>0.107</td>
</tr>
</tbody>
</table>

1 Elasticities calculated using the expression: $e_i = (x_{ij}/\hat{P}_i)\left(\frac{\partial \hat{P}_i}{\partial x_{ij}}\right)$ for each Variety i and variable j.

Table 2. Category Choice Model Response Elasticities by Segment

<table>
<thead>
<tr>
<th>Variable:</th>
<th>Segment 1</th>
<th>Segment 2</th>
<th>t-ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>LP$^1$</td>
<td>0.972*</td>
<td>0.917</td>
<td>16.925</td>
</tr>
<tr>
<td>ALOY</td>
<td>0.090*</td>
<td>0.039</td>
<td>10.187</td>
</tr>
<tr>
<td>CR</td>
<td>0.142*</td>
<td>0.091</td>
<td>-3.223</td>
</tr>
<tr>
<td>INV</td>
<td>-0.015*</td>
<td>-0.022</td>
<td>17.553</td>
</tr>
<tr>
<td>AAD</td>
<td>0.450*</td>
<td>0.175</td>
<td>142.006</td>
</tr>
<tr>
<td>OAD</td>
<td>-0.069*</td>
<td>-0.066</td>
<td>-3.090</td>
</tr>
</tbody>
</table>

1 A single asterisk indicates mean elasticities are significantly different from each other at a 5% level of significance.