MULTIVARIATE CLUSTER ANALYTIC MODEL OF BENEFIT-BASED MARKET SEGMENTATION: A CASE STUDY FROM THE RECREATION AND LEISURE INDUSTRY

Dana A. VanDen Heuvel, St. Norbert College
P. Raj Devasagayam, St. Norbert College

ABSTRACT

Past research in benefit-based market segmentation has predominantly relied on conjoint analytic models as methodological tools. The authors of this research underscore the flaws of conjoint analytic models and propose a cluster analytic model as a possible solution. Further, a market in the leisure and recreation industry is segmented via the use of a computer-based segmentation procedure using multivariate cluster analysis. The model was used to segment the markets of a local zoo in the Midwest. The consequent market segments were clearly identifiable and valid. The segmentation scheme has proved useful in strategic planning initiatives that have enhanced customer visitation and satisfaction at the zoo.

INTRODUCTION

Increased competition has rendered the tasks of market segmentation, market targeting and product positioning of paramount importance. Companies may gain sustainable competitive advantage through innovative segmentation of their markets. Distinct market segmentation provides a vehicle for directing one=s marketing efforts toward that narrower base of potential customers. The trick to marketing to a narrower base of customers is figuring out which narrow portion to market to. A market segmentation procedure that achieves such precision, yet has received scarce attention from marketing scholars is the benefit-based market segmentation schema.

Benefit-based market segmentation involves segmenting the market for a product based on the intrinsic value that customers derive from the product. This intrinsic value (benefit) could take various forms such as snob appeal, perceived quality, or derived economy. This method of market segmentation is significantly different from other methods of market segmentation. It seeks to get "inside" the consumer=s cognitive processes and find out what the customer is feeling about our product, rather than arbitrarily identifying a consumer based on how much money she makes or where he lives. The underlying principle is that consumers do not seek a product per se, rather they seek what that product does for them, or the derived benefit. With benefit-based segmentation, the objective is to identify *why* the customer is buying the product, and group customers with similar *whys* together.

PRIOR RESEARCH

Past research has shown that the benefits consumers derive from various product and service attributes are powerful discriminatory variables for market segmentation (Haley 1968; Wind 1978). Traditionally marketing scholars have restricted benefit-based market segmentation schema to conjoint analytic models (DeSarbo, Oliver, and Rangaswamy 1989; De Soete and DeSarbo 1991; Green and Krieger 1991; Kamakura 1988; Wittink, Vriens, and Burhenne 1994). There are several two-stage conjoint models in use to facilitate benefit-based market segmentation. However, the large numbers of such models merely confuse and frustrate both researchers and practitioners. Consequently, conjoint analytic segmentation models are largely chosen based on availability of expertise and software, coincidental acquaintance with scholars or algorithms, or personal preferences. The concern for strengths and weaknesses of models of choice and the propriety of use of a specific model for a specific application are often ignored.

One can't help but notice the irony of picking an analytic tool to segment market by derived benefit based on factors *other* than product attributes and true benefits!

This study provides an alternative to conjoint analytic models of benefit-based market segmentation, namely a multivariate cluster analytic model. We will soon see that this model is easy to comprehend, relatively effortless in application, and involves the managers responsible for product strategy formulation in interpretation of results. The cooperative efforts of managers and researchers strengthen the validity of results and findings.

A cluster analytic model has several methodological advantages. Firstly, the advantage of not depending on unreliable individual-level estimates of benefits based on any subjective *a priori* segmentation scheme, as often seen in conjoint analytic models founded on Hagerty's (1985) two-stage procedures. Secondly, conjoint models often suffer from insufficient degrees of freedom that may lead to incorrect partworth estimates and misclassification of subjects (Kamakura 1988; Umesh and Mishra 1990). As we will see, a cluster analytic model allows for ample degrees of freedom provided the sample size is adequate. Thirdly, cluster analytic models don't face identification problems often witnessed in alternative models that overparametrize degenerate solution. A related issue thus involves respondent fatigue associated with conjoint specific data collection procedures. Often the task of attribute weighting, attribute-level determination, and preference statements and ratings can be foreboding for most respondents. In a cluster analytic model interval level data is obtained with relatively negligible respondent burden, further strengthening the validity of results obtained.

THEORETICAL RATIONALE

A cluster analytic model of benefit-based segmentation is concerned with devising a methodology of segmenting a very large and diverse potential market, through the use of a statistical algorithm, with the application of very specific benefit-oriented finite variables. In the search for narrow and specific segments there are three criteria that need to be satisfied, homogeneity within the segment, heterogeneity across the segments, and aggregation throughout segments.

Homogeneity within the segment is critical to defining the target customer. In the case of a benefit-based segmentation scheme, customers that derive the most similar benefits would be grouped in one segment. For instance, in the leisure and recreation industry it is a common practice to group customers who have young children and those that are on a tight budget. In this case, the common desired benefit being sought is economy. Heterogeneity across the segments allows for the differentiation of segments and consumers. If two customers have very different needs, such as economy and snob appeal, they would be grouped in two very heterogeneous segments. By making the segments as heterogeneous as possible, the company ensures that marketing resources are utilized efficiently. Finally, aggregation is the process by which customers are grouped together. Segmentation was once thought of as being a desegregation of customers, on the contrary, the objective is to group them together for the most effective targeting, not separate them out to be targeted separately. The process of aggregation is expected to result in market segments that satisfy the aforementioned twin qualifications of homogeneity within and heterogeneity across segments.

METHODOLOGICAL RATIONALE AND A CASE STUDY

The rationale behind the use of cluster analysis ties in very closely to the rationale for market segmentation. Cluster analysis seeks to organize information about variables so that relatively homogeneous groups, or clusters, can be formed (Anderson 1984; Dillon and Goldstein 1984; Lance and Williams 1967). Similar to market segmentation, the clusters that are formed should be highly internally homogeneous as well as externally heterogeneous as the groups should be as dissimilar as possible. Cluster analysis, like market segmentation, is also a gathering process, since responses are grouped together based on likeness of variables (Joreskog 1977). Cluster analysis performs agglomerative hierarchical clustering of data in order to classify observations into *prima fascia* unknown groups. Observations are aggregated in a stepwise fashion. Each successive step reduces the number of clusters by one. Similarity and distances are computed based on the scale on which distances are defined and the methods of linking researchers prescribe (Morrison 1976).

In this research, a market in the leisure and recreation industry is segmented via the use of a computer-based segmentation procedure using multivariate cluster analysis. Data was collected from a local zoo in the Midwest and

cluster analysis was used to create a multivariate model to achieve a very precise grouping of responses. Findings indicate that the resultant market segments were clearly identifiable and valid. Results obtained from the study have resulted in strategic planning initiatives enhancing customer visitation and satisfaction.

A preliminary draft of the data collection instrument was pre-tested with undergraduate students and the zoo management. The final product turned out to be a rather well formatted and concise questionnaire consisting of 26 items covering a range of topics that pertained to the zoo and strategic issues facing the management. The data was procured from a convenience-based random sample of visitors at the local zoo. Personal interviews in a mall-intercept fashion were conducted. A high response rate was made possible through personal interviews assuring respondents of anonymity, stating the contribution of the study in improving their future visits to the zoo, and offering a token incentive for participation in the study. A total of 743 surveys were returned out of a possible 1100 (Response Rate 67.55%).

Descriptive statistics on the sample profile indicate that 56.28% of the people visiting the zoo were female, while 43.44% were male. The largest number (38.16%) of visitors fell between the ages of 25-39. Twenty-two percent of the visitors make between $30,000 and $39,999 annually. The majority of visitors, 51.98% are married, while 37.71% reported being single. People with children accounted for 57.69% of all visitors, 36.14% had two children (under 18) in their home. 62.76% of the visitors reside within the county and 45.11% traveled less than 15 miles to get the zoo. The typical visitor is a married female, with 2 children, has an income between $30,000 and $39,999, and travels less than 15 miles to get to the zoo.

TABLE 1

Segments	Number of Observations	Within Cluster Sum of Squares
Cluster 1	20	0.000
Cluster 2	24	0.000
Cluster 3	38	0.000
Cluster 4	25	0.000
Cluster 5	28	0.000
Cluster 6	1	0.000
Cluster 7	22	0.000
Cluster 8	25	0.000
Cluster 9	11	0.000
Cluster 10	50	0.000
Cluster 11	10	0.000
Cluster 12	31	0.000
Cluster 13	20	0.000
Cluster 14	35	0.000
Cluster 15	20	0.000
Cluster 16	12	0.000
Cluster 17	8	0.875
Cluster 18	27	0.000
Cluster 19	138	206.174
Cluster 20	176	216.875

From this survey, we selected several questions that depicted benefits that the zoo visitors might derive from the zoo. These selections ultimately became the variables that were used in the cluster analysis. One of the issues facing the zoo was to ascertain the profitability and feasibility of introducing learning programs on animals, environment, and

coexistence at the zoo. The concern was to determine if there was a sizable portion of the market that sought the benefit of viewing the zoo as an interactive educational institution for their children in addition to the benefit of its being a zoological preserve. In order to aid the decision making process, three variables that dealt with income levels, number of children in the household, and willingness to support a learning program at the zoo were analyzed using the multivariate cluster analytic model. After the cluster analysis was run, a graph or Dendrogram was produced that graphically depicted the clusters that we were seeking. The next step is often referred to as 'cutting the Dendrogram' (Johnson and Wichern 1988; Milligan 1980). Clusters are identified and *cut* based on a Cut or Amalgamation Statistic provided by the procedure and researcher judgement (Morrison 1976).

In this particular case 20 distinct market segments emerged. Based on within cluster sum of squares the most homogeneous segments were identified (Table 1). We were seeking within cluster sum of squares to be as close to zero as possible indicating a tight spread around the centroid. In addition we were interested in clusters being as dissimilar to each other as possible in order to achieve heterogeneity across segments. We use the distances between cluster centroids to ensure heterogeneity. Although we now have very distinct clusters, or segments, the task of segmenting the market is far from complete. The final steps that need to be taken involve the application of the initial variables to the clusters and the decision making process about what the clusters actually mean. Labels accorded to each segment are largely based on inputs obtained from zoo management. Eventually, based on management input two feasible segments emerge which seek the benefits offered by a learning program at the zoo.

LIMITATIONS AND FUTURE RESEARCH

During the course of implementing this research project several limitations became apparent. First, our data was obtained late in the season (in the face of severe winters in the local area) and as a result zoo visitation numbers were lower than normal. Sampling errors might have crept in. Second, there were no past studies available to compare our results across samples or alternate segmentation models.

Use of a cluster analytic model imposes certain limitations on our findings. First, the data set that is utilized when performing cluster analysis needs to be one of perfection (Milligan 1980). There is no provision in the cluster analysis procedure for missing values. Secondly, the need for superior computing power arises when conducting this type of analysis.

The lack of a direct comparison between conjoint and cluster analysis is a weakness that was unavoidable in our study. It was not possible to procure the data necessary to perform a conjoint analysis and the cluster analysis without angering our respondent base. Duplication of information throughout the sample would have been necessary in order to overcome validity issues raised by comparing models across two samples. One would always doubt if results were indeed dissimilar due to differences in models or because of differences in sample. The need for specialized data to run conjoint-based models further illustrate the ease with which multivariate cluster analysis can be performed with rather simple data from a standardized and versatile questionnaire.

In summary, we have presented a statistical model using multivariate cluster analysis as a possible alternative to the error prone conjoint model of benefit-based market segmentation. This model, if proven effective, could be applied to other businesses and industries. In the instance of the case study that was performed several distinctive clusters were obtained. These clusters exhibited the traits that are desired in market segments B homogeneity within the segment and heterogeneity across the segments. However, given the very infantile nature of this type of statistical theory, there is much work to be done in getting to the point where multivariate cluster analysis is a reliable and accepted method for producing benefit-based market segments.

REFERENCES

Anderson, T. W. (1984), "An Introduction to Multivariate Statistical Analysis," Second Edition, John Wiley & Sons.

DeSarbo, Wayne S., Richard L. Oliver, and Arvind Rangaswamy (1989), "A Simulated Annealing Methodology and Clusterwise Linear Regression," Psychometrika, 54 (4), 707-36.

De Soete, Geert and Wayne S. DeSarbo (1991), "A Latent Class Probit Model for Analyzing Pick Any/N Data," Journal of Classification, 8 (7), 45-63.

Dillon, William and M. Goldstein (1984), "Multivariate Analysis: Methods and Applications," John Wiley & Sons.

Green, Paul E. and Abba M. Krieger (1991), " Segmenting Markets with Conjoint Analysis," Journal of Marketing, 55 (October), 20-31.

Hagerty, Michael R. (1985), "Improving the Predictive Power of Conjoint Analysis: The Use of Factor Analysis AND Cluster Analysis," Journal of Marketing Research, 22 (May), 168-84.

Haley, Russel I. (1968), "Benefit Segmentation: A Decision-Oriented Research Tool," Journal of Marketing, 32 (July), 30-35.

Harmon, H. (1976), "Modern Factor Analysis," Third Edition, University of Chicago Press.

Johnson, R. and D. Wichern (1988), "Applied Multivariate Statistical Methods," Second Edition, Prentice Hall.

Joreskog, K. (1977), "Factor Analysis by Least Squares and Maximum Likelihood Methods," Statistical Methods for Digital Computers, Ed. K. Enslein, A. Ralston and H. Wilf, John Wiley & Sons.

Kamakura, Wagner A. (1988), "A Least Squares Procedure for Benefit Segmentation with Conjoint Experiments," Journal of Marketing Research, 25 (May), 157-67.

Lance, G. N. and W. T. Williams (1967), "A General Theory of Classificatory Sorting Strategies, I. Hierarchical Systems," Computer Journal, 9, 373-380.

Milligan, G. W. (1980), "An Examination of the Effects of Six Types of Error Perturbation on Fifteen Clustering Algorithms," Psychometrika, 45, 325-342.

Morrison, D. (1976), "Multivariate Statistical Methods," McGraw-Hill.

Umesh, U. N. and Sanjay Mishra (1990), "A Monte Carlo Investigation of Conjoint Analysis Index-of-Fit: Significance and Power," Psychometrika, 55 (1), 33-44.

Wind, Yoram (1978), "Issues and Advances in Segmentation Research," Journal of Marketing Research, 15 (August), 317-337.

Wittnik, Dick R., Marco Vriens, and Wim Burhenne (1994), "Commercial Use of Conjoint Analysis in Europe, Results and Critical Reflections," International Journal of Research in Marketing, 11 (1), 41-52.

-***-