ESTIMATION OF MEXICAN MARKET SEGMENTS (MEMS) COMPARISON OF ALTERNATIVE STRATEGIES FOR SEGMENT DEFINITION

Sarah Maxwell, Florida International University

Introduction

In 1991, the Center for International Business and Environment Studies (CIBES) surveyed the Mexican market, making available appropriate data for the present study of the Mexican Market Segments (MEMS). This study followed the customary technique for defining market segments: that is to cluster a battery of psychographic variables. The goal is the recovery of clearly defined, distinctly separated clusters. However, this goal can be hampered by "noisy" variables which do not contribute to, and may actually mask, the underlying cluster structure.

The purpose of the MEMS's research, therefore, was to compare alternative strategies of clustering in an effort to recover not just *adequate* but *optimal* segments of the Mexican market. The variables were clustered using three different analysis strategies: the traditional methods of clustering (a) "raw" data and (b) factor scores plus (c) a new method of preprocessing "raw" data points with Heuristic Identification of Noisy Variables (HINoV).

The optimality of the three strategies was evaluated by (1) *predictive validity* as measured with the "confusion index" of discriminant analysis, (2) *external validity* as measured by the usefulness and meaningfulness of the demographic profiles of each cluster, and (3) *cross validation* as measured by the stability of profiles across split-half subsets.

Market Segmentation

Wendell Smith's article (1956) in the *Journal of Marketing* is often cited as the introduction of market segmentation. But Smith himself (1978) credited Wroe Alderson for having established the concept a quarter century before. Certainly by 1978, when Yoram Wind edited the now classic status report on segmentation research in the *Journal of Marketing Research*, all the bases of segmentation were in place.

The major work in market segmentation has been in the applied area. Marketing practitioners have used segmentation to target specific markets, position products, and predict group responses to new products or to changes in the marketing mix. In industry, two basic segmentation methods have been utilized (Wind 1978): *a priori* segmentation using geographic and demographic measures (age, sex, income, education, occupation, etc.) and *post hoc* segmentation using a battery of interrelated variables to create segments of necessarily unknown number and size. These variables relate to the respondents' activities, interests, and opinions and reflect their life style or psychographic profile.

Segmenting the market with psychographics has become increasingly popular (Wind 1978, Michman 1991). Psychographic segmentation assumes that similar purchasing motivations stem from a similar human value system. Values are useful for market segmentation because they have been shown to be a primary determinant of purchasing behavior and stable over time (Henry 1976). Rokeach (1973) defined values as "standards of desirability invoked in social interaction to evaluate the preferability of behavioral goals or modes of action."

Many different schemes of psychographic segmentation were spawned by the Rokeach model. The two most prominent were VALS ("Values and Life Styles") and LOV ("List of Values"). Since VALS is proprietary information, the method of classifying groups is not published. In both the Rokeach and LOV models, the groups are first determined by clustering. Individuals are then classified into the pre-specified cluster group simply on the basis of their top-ranked value. (Kamakura & Mazzon 1991). The predictive utility of the groupings is then evaluated with regression.

Clustering Strategies

While the framework of psychographic segmentation appears to have gained general acceptance, there is still considerable disagreement over the methods. The debate over clustering strategies concerns the issue of pre-processing the data or dealing directly with the raw data.

The first strategy implemented in the MEMS's study was the traditional clustering of the raw data. The clustering technique utilized was the same as for the other two strategies: namely, an iterative partitioning algorithm as opposed to a hierarchical technique. Iterative partitioning, specifically the K-means partitioning algorithm, has gained acceptance as the preferred method for handling the large data sets typical in marketing (Green, Carmone & Kim 1990). The K-means algorithm starts from K pre-determined cluster centers. Cases are assigned to the cluster with the nearest centroid. The centroid is then recalculated and cases reassigned until the variance within each cluster is minimized.

The second traditional strategy implemented in the MEMS's study was the clustering of factor scores as had been recommended by various researchers (Wind 1978; Punj & Stewart 1983). The value of factor analysis is that it removes the redundancy among the underlying attributes and reduces the number of variables.

The third strategy implemented was HINoV, a new strategy which has been developed in response to a problem inherent in both raw data and factor score clustering: the presence of outliers and noisy variables which can wreak havoc with the definition of clusters. Even one irrelevant variable can distort the cluster solution (Punj & Stewart 1983) and mask the underlying cluster structure (DeSarbo, Carroll, Clark & Green 1984). Irrelevant or "noisy" variables need to be identified and removed. As Green & Kim (1988) conclude, "the problem of variable selection remains of major importance in cluster analysis."

HINoV was developed to address this problem (Carmone 1991). It makes a pairwise comparison of cluster membership of every variable with that of every other variable. The amount of agreement between each pairwise comparison is measured with an adjusted Rand Index. The Rand Index gives the ratio (between 0 and 1) of agreement of cluster membership to disagreement. If two clustering solutions contain the exact same members, the Rand index would be 1.0. If they contain none of the same members, the Rand Index would be 0.0. The Index has been modified (Hubert & Arabie 1985) to remove the bias due to chance. This adjustment was documented as superior in the studies conducted by Milligan (1989). The Rand Indices for all pairwise comparisons are summed for each variable. This indicates how close that particular variable's clustering solution is to the clustering solutions of the other variables. It measures the contribution of that individual variable to the overall clustering solution. These summed Rand Indices are then ranked and plotted to give a visual indication of which variables can be removed.

Clustering Evaluation

Evaluating whether the results from one clustering strategy are "better" than those from another is difficult on at least two counts: (1) Clustering itself has been called a "fuzzy" concept, making some researchers conclude that clusters may not exist (Tryon & Bailey 1970). (2) Clustering, unlike other multivariate techniques, lacks statistical procedures for establishing external reliability and validity (Punj & Stewart 1983).

Monte Carlo studies using simulated data have been used to evaluate various clustering algorithms, preprocessing procedures, etc. (Milligan 1985, Kara 1991). However, Monte Carlo methods cannot be used with real data where the structure is unknown. For the MEMS's study, therefore, surrogates had to be used to evaluate predictive, external,and cross validation.

1. Predictive Validity: Several studies have recommended that discriminant analysis be used to evaluate predictive validity (Punj & Stewart 1983). Cluster membership is the dependent variable with activities, interests, and opinions as the predictor variables (Kamakura & Mazzon 1991). A randomly selected hold-out sample is used to test the model developed from the original sample (minus the hold-out). The percentage correctly predicted is an indication of the model's validity.

The percentage is expected to be generally high since the clusters and the discriminant model are determined from the same variables. However, in the MEMS's study a *comparatively* higher percentage correct was used as an indication that one clustering solution was "better" than another.

2. External Validity: In clustering, the term "external validity" is used in a more applied manner than in formal theory evaluation. Rather than referring to the generalizability of a causal relationship, it refers to how meaningful the information is to the end user. As Punj & Stewart point out (1983), "external validity requires a demonstration that the clusters are useful in some larger sense." The end user needs market segments which are identifiable, understandable, and actionable. In the MEMS's study, market segments which met this criterion were therefore considered "better" than those which did not.

3. Cross Validation: Stability across split sets of data is often used as a surrogate for the concept of reliability in classical test theory (Punj & Stewart 1983). Clustering is carried out on each randomly selected half of the data set. Descriptive statistics of each set are then compared to determine whether the segments hold up. When profiles across sets in the MEMS's study were more similar, the clustering solution was considered "better".

Methodology

The data source for the MEMS's study, CIBES, was founded at Florida International University in 1991 to provide academic and business-related marketing research in Latin America. The major emphasis for CIBES's initial, ground-breaking project was to establish an Index of Consumer Confidence in Mexico. To this end, personal interviews were conducted with a probability sample of 900 households in six geographic areas of Mexico. The data were collected using a questionnaire with a Likert-like 5-point scale. The instrument was developed and carefully pre-tested by the CIBES's staff.

The CIBES's survey questions related to psychographics, demographics, consumers' confidence, plus client-oriented and consumption-oriented proprietary sections pertaining to products of sponsoring companies. The MEMS's study used two subsets of the CIBES's data: 60 psychographic variables plus 32 client variables used to verify results In addition, 23 demographic variables were used to build segment profiles.

Prior to the MEMS's study, the data had been cleaned by the CIBES's staff to correct for input and/or coding errors. Missing data constituted only .3% of the responses with no more than .7% missing in any one variable. Because the clustering algorithms required complete data, the missing data were recoded to the rounded integer value of each variable's mean response. To identify "suspicious" interviewers or respondents, the data sets were transposed and the frequency of responses by respondent was counted. Respondents thought to be uninterested or unthinking because they answered all the questions with the same response were identified and discarded. This resulted in the removal of 10 psychographic cases and four client cases.

Preliminary analysis was done to determine the optimal number of clusters, an evaluation which is "more an art than a science" (Kamakura & Mazzon 1991). The MEMS's study consequently applied two different methods. First, the Wilks' Lambda and F-Value of each cluster solution from 2 to 12 clusters for the psychographic and the client data sets were calculated using MANOVA. Four subsets of each data set were then analyzed with hierarchical clustering to determine a "fusion index". Although the results tended to favor a four-cluster solution, they were not conclusive. Consequently, in comparing factor score vs. raw data clustering, all solutions for 2 to 12 clusters were reviewed.

Each of the two data sets was separated into two equal odd/even sets of 450 respondents each. Since the sequence of responses followed the geography of the interviews, an odd/even bifurcation was considered a legitimate means of maintaining the randomness of the sample. The split-half sets of both the psychographic and client data sets were analyzed by first raw data clustering and then factor score clustering for two to twelve cluster solutions.

The SPSS QUICK CLUSTER program was used to cluster both the raw data and the factor scores by the K-means iterative partitioning algorithm. The program selects well-spaced cases as initial centers or "seeds". In the MEMS's study the centroids of the solution were then used to "seed" a subsequent iteration of the program twice more as recommended by SPSS. Factor analysis was done with principal components and varimax rotation. Factor scores were determined via regression and used as the input for the clustering procedure.

1.Predictive validity was evaluated by using cluster membership as the dependent variable in discriminant analysis. 50% of the data (225 cases in each set) was withheld. The model was built from the initial sub-sample and then used to predict group membership of the hold-out sample. The "confusion matrix" compared the percentage of cases correctly classified.

2.External validity: The managerial usefulness of the solutions was measured by cross-tabbing group membership with demographic variables and finding how many demographic variables were significant. The assumption was that a greater number of significant differences in the clusters was potentially more useful to a manager than a fewer number.

3. Cross Validation was evaluated by comparing the cluster means and standard deviation for each of the key variables.

The comparison of HINoV and raw data clustering followed basically the same steps as the comparison of raw data clustering and factor score clustering. The differences were an additional cleaning of low variance cases and variables plus the limitation of the analysis to a four-cluster solution. This was a necessary limitation since HINoV is computer-time intensive. A four-cluster solution was selected based on the earlier, albeit inconclusive, evidence in its favor. As a preliminary step, persistent outliers were removed. This meant that a total of 16 cases were removed from the psychographic data set and 13 cases from the client data set.

Conventional methods were initially used to identify the variables not contributing to the solution. Those having a disproportionately high number of responses in two adjacent values were removed. As a result, 10 psychographic variables and two client variables were discarded. The HINoV procedure was applied as an additional means to identify "noisy" variables. The two data sets and the split-half sets of each were then evaluated.

Results

The first-stage comparisons of the 88 raw data vs. factor score clustering solutions gave mixed results. The second-stage comparisons of the four-cluster solutions for raw data, factor score and HINoV gave results which favored HINoV.

Raw Data vs. Factor Score Clustering:

1. Predictive Validity: The results from the discriminant analysis predictions of cluster membership were conflicting: raw data clustering gave better results with the client data set and factor score clustering gave better results with the psychographic. These results may have reflected the fact that the larger (60-variable as opposed to 32-variable) psychographic data set had greater redundancy which factor analysis removed. In any case, the results did not clearly indicate the superiority of raw data clustering or factor score clustering.

2. External Validity: The analysis of the crosstabs for different number of clusters showed that with raw data clustering, 41% of the crosstabs between cluster membership and 23 demographic variables were significant. With factor score clustering, only 15% were significant. This suggested that raw data clustering provided more useful information to a manager.

3. Cross Validation: A comparison of the significant crosstabs in the split-half samples of each data set showed greater potential similarity of profiles with the raw data clustering than with the clustering of factor scores. With raw data clustering, three times as many crosstabs were significant in both the odd and even split half samples of the client and psychographic data.

HINoV vs. Raw Data & Factor Score Clustering

1.Predictive Validity: A higher percentage of group membership was correctly predicted by discriminant analysis for HINoV clustering than with raw data or factor score clustering. The HINoV results, based on only one-third (21) of the 60 psychographic variables, predicted cluster membership more accurately than either of the alternatives. The results were substantiated by the high average score of the split run and further supported with the client data set.

2.External Validity: The profiles of the psychographic clusters built on the HINoV selected variables appeared to be more clearly differentiated than those built on raw data clustering. With raw data clustering, the first two groups identified were virtually identical in their demographic pr ofiles; they differed only in their attitudes. The first group was less happy, less religious, less family oriented. While this was an interesting distinction, it would not be managerially useful because it would not be actionable. The HINoV clustering appeared to be more meaningful, defining four distinctly different groups.

3. Cross Validation: A comparison of the profiles for the odd/even data sub-samples was made to determine the stability of the results with HINoV. Too few of the psychographic/demographic crosstabs were significant to make a comparison. The crosstabs of the client groups and their demographics, however, showed nine significant demographic variables. With four clusters, this meant 36 profile characteristics, of which 28 (80%) were identical. Another four characteristics were similar with one group having an "average" response which might have been in the same general direction as the second group but not as strong. These results indicated relative stability of the HINoV solution.

Psychographic Groups Correctly Predicted

STRATEGY	CASES RE-MOVED	VARI-ABLES REMOVED	% CORRECT
RAW DATA (AVG)	10	0	77%
FACTOR SCORE(AVG)	10	0	83%
RAW DATA	13	0	82%
RAW DATA	13	10	78%
HINoV	13	29	87%
HINoV	13	39	92%
HINoV (AVG SPLIT)	13	39	91%

Conclusion

This study demonstrated that optimal psychographic segments for MEMS were recovered with the HINoV strategy. The "noisy" variables distorting the solution were apparently identified and removed. The remaining variables were those which contributed most to the clustering solutions. The resulting clusters were "better" in the sense that they were higher in predictive and external validity plus cross validation. The removal of HINoV identified variables helped define more understandable and clearly differentiated segments. Furthermore, HINoV improved the validity of the clusters using only one-third to one-half of the variables.

As an immediate benefit, clear segments of the Mexican market have been identified, a valuable contribution to future marketing in Mexico under the North American Free Trade Agreement. As a long term benefit in ongoing projects such as the study of the Mexican market by

Client Groups Correctly Predicted			
STRATEGY	CASES REMOVED	VARIABLES REMOVED	% COR-RECT
RAW DATA (AVG)	4	0	82%
FACTOR SC.(AVG)	4	0	77%
RAW DATA	16	0	92%
RAW DATA	16	2	87%
HINoV	16	14	93%
HINoV (AVG SPLIT)	16	14	86%

CIBES, being able to reduce the number of psychographic questions can reduce the cost of the survey as well as the fatigue of the respondent.

References

Blashfield, Roger K., and Mark S. Aldenderfer. 1978. "The Literature on Cluster Analysis" *Multivariate Behavioral Research* 13:2 271-295.

Carmone, Frank J. Jr. 1991. "An Algorithm for Heuristically Identifying Noisy Variables in Cluster Analysis." *Unpublished Working Paper.*

DeSarbo, Wayne S., and J. Douglas Carroll, Linda A. Clark, Paul E. Green. 1984. "Synthesized Clustering: A Method for Amalgamating Alternative Clustering Bases with Differential Weighting of Variables." *Psychometrika* 49:1 57-78.

Green, Paul, Abba Krieger and Catherine M. Schaffer. 1985. "Quick and Simple Benefit Segmentation" *Journal of Advertising Research.* 25:3 9-18.

Green, Paul, and Jonathan Kim. 1988. "Optimal Variable Weighting Methods in Cluster Analysis: How well do they Validate? " *Unpublished Working Paper.*

Green, Paul, Frank Carmone, and J. Kim. 1990. "A Preliminary Study of Optimal Variable Weighting in K-Means Clustering" *Journal of Classification* 7:2 271-286.

Henry, Walter A. 1976. "Cultural Values Do Correlate with Consumer Behavior." *Journal of Marketing Research.* 13:2 121-127.

Hubert, Lawrence, and P. Arabie. 1985. "Comparing Partitions." *Journal of Classification.* 2:2 193-218.

Kamakura, Wagner A., and Jose A Mazzon. 1991. "Values Segmentation: A Model for the Measurement of Values and Value Systems." *Journal of Consumer Research.* 18:2 208-218.

Kara, Ali. 1991. "An Empirical Investigation of the Robustness of 'Heuristic Identification of Noise Variables (HINoV)'." *Unpublished Working Paper*

Michman, Ronald D. 1991. *Lifestyle Market Segmentation.* New York:Praeger.

Milligan, Glenn W. 1989. "A Study of a Variable Weighting Algorithm for Cluster Analysis." *Journal of Classification.* 6:1 53-71.

Punj, Girish, and David W. Stewart. 1983. "Cluster Analysis in Marketing Research: Review & Suggestions for Application" *Journal of Marketing Research.* 20:2 134-148.

Rokeach, Milton. 1973. *The Nature of Human Values.* New York, NY: The Free Press, Macmillian Publishing Co.

Smith, Wendell R. 1956. "Product Differentiation and Market Segmentation as Alternative Marketing Strategies." *Journal of Marketing.* 21:7 3-8.

Smith, Wendell R. 1978. "Retrospective Note on Market Segmentatin." *Journal of Marketing Research.* 15:3 316.

Tryon, Robert C., and Daniel E. Bailey. 1970. *Cluster Analysis.* New York, NY: McGraw Hill.

Wind, Yoram. 1978. "Issues and Advances in Segmentation Research." *Journal of Marketing Research.* 15:3 317-337.