Market Segmentation with Choice-Based Conjoint Analysis

WAYNE S. DESARBO
School of Business Administration, University of Michigan, Ann Arbor, MI 48109-1234

VENKATRAM RAMASWAMY
School of Business Administration, University of Michigan

STEVEN H. COHEN
Stratford Associates, Boston, MA

Abstract

Choice-based conjoint analysis has increased in popularity in recent years among marketing practitioners. The typical practice is to estimate choice-based conjoint models at the aggregate level, given insufficient data for individual-level estimation of part-worths. We discuss a method for market segmentation with choice-based conjoint models. This method determines the number of market segments, the size of each market segment, and the values of segment-level conjoint part-worths using commonly collected conjoint choice data. A major advantage of the proposed method is that current (incomplete) data collection approaches for choice-based conjoint analysis can still be used for market segmentation without having to collect additional data. We illustrate the proposed method using commercial conjoint choice data gathered in a new concept test for a major consumer packaged goods company. We also compare the proposed method with an *a priori* segmentation approach based on individual choice frequencies.

1. Introduction

In recent years, choice-based conjoint analysis (Louviere and Woodworth, 1983) has increased in popularity in the marketing research community (Struhl, 1994). Choice-based conjoint analysis has several advantages. First, data collection involves simulated purchase decisions (choices), which is a more realistic and relatively simple task for respondents than are ratings or rankings. Second, the derived part-worth utilities reflect impact on product choice rather than a change in rating or ranking. Hence, share estimates are directly estimated instead of requiring a two-stage process (estimation followed by choice simulation) that requires many assumptions (such as choice rules). Third, product-specific attributes or levels (unique features of a brand, different price levels for each brand) can be easily accommodated and brand-specific utilities can be estimated. Fourth, in market forecasting applications where more accurate estimates of share and demand are desired, choice-based conjoint analysis has the ability to embed a no-choice option. Fifth, unlike traditional conjoint analysis, there is more flexibility in designs encompassing a wide range of choice contexts such as choosing among alternate product types (asking respondents to choose among different types of portable communication devices), competing products within the

same brand (different telephone calling plans), line extensions and product variants (as illustrated in our empirical application), and so on.

Market segmentation has been a primary objective in performing traditional (ratings-based) conjoint analysis in commercial applications (Wittink and Cattin, 1989). Green and Krieger (1991) indicate that a *post hoc* approach to market segmentation is commonly employed with traditional conjoint analysis wherein estimated individual-level part-worths are clustered to obtain buyer segments with similar preferences and price sensitivities. In the case of choice-based conjoint analysis, it is infeasible to obtain individual-level parameter estimates due to the lack of adequate observations per respondent. Hence, unlike ratings based conjoint analysis, it is impractical to form market segments *post hoc* by clustering individual-level parameter estimates (Elrod, Louviere, and Davey, 1992). The typical current practice is to conduct aggregate-level estimation (Struhl, 1994). As Struhl (1994, p. 38) notes: "Aggregate-level analysis means no individual-level utilities of the type obtained from conjoint. And the absence of individual-level utilities means segmentation is not available. ... Segmentation based on conjoint utilities often provides tremendously useful results. The groups that emerge usually have sharply different wants and needs related to the product in question." While one can still perform *a priori* segmentation using ancillary variables (such as demographics) or based on the observed choice frequencies, there is no guarantee that the delineated segments will exhibit different preferences and price sensitivities (Green and Krieger, 1991).

In this paper, we discuss a method for *post hoc* market segmentation (based on the part-worths) with choice-based conjoint data. The proposed method is a generalization of the Kamakura and Russell (1989) probabilistic segmentation approach to an experimental context entailing choice data gathered over (incomplete) conjoint choice sets. Using this method, we demonstrate how one can simultaneously perform market segmentation and calibrate the segment-level part-worths without collecting any additional data. The number of segments, sizes of each segment, and values of segment-level part-worths are all obtained using only the observed conjoint choice data. Conceptually, the process is similar to DeSarbo, Wedel, Vriens, and Ramaswamy (1992), who conduct market segmentation and segment-level part-worth estimation simultaneously in the context of ratings-based conjoint analysis. However, their approach cannot be applied to choice-based conjoint analysis, which involves the collection of *choice* data over (incomplete) fractionated choice set designs (that is, each respondent need not see all product or attribute profiles in a fractionated design) and the specification of a discrete choice model (such as logit model) for parameter estimation. Given the aforementioned appeal of choice-based conjoint models with practitioners, the proposed method extends the domain of applicability of these models to accommodate market segmentation applications.

2. Method

Let

$i = 1, \ldots, I$ respondents;
$j = 1, \ldots, J$ conjoint profiles and brands;
$k = 1, \ldots, K$ conjoint attributes and dummy variables;

$n = 1, \ldots, N$ choice sets (such as from a 2^J design, Louviere and Woodworth, 1983);

$C_n =$ the specific brands in the nth choice set;

$X_{jk} = k$th dummy variable for the jth conjoint profile;

$s = 1, \ldots, S$ market segments;

$\beta_{ks} =$ the impact coefficient for the kth attribute for market segment s;

$Y_{ijn} = 1$ if respondent i chooses brand j in the nth choice set among C_n; 0 otherwise;

In a typical choice-based conjoint experiment, a respondent is asked to choose an alternative from a competitive set of alternatives or choice set, each of which is a profile of different combination of levels or values of a set of multiple attributes. The respondent repeats this task for a limited number of choice sets thereby providing repeated choice data. Louviere and Woodworth (1983) discuss different approaches to the design and generation of these choices sets. These include fixed and 2^J choice set designs. In a fixed-choice set design, the number of alternatives (e.g., brands) in the choice set is fixed over the respondent's task; in other words, there are only changes in one or more attributes of a fixed number of alternatives. In a 2^J design, the choice set size can vary over the respondent's task as well, although a constant base is required (such as a no-choice option or some reference product) for estimation purposes.

Louviere and Woodworth (1983) discuss in detail the implementation of such 2^J designs. Based on our experience, the fixed-choice set design is relatively more popular among practitioners due to its simplicity. Nevertheless, in the development of our approach, we assume the more general case of a 2^J design; the fixed-choice set design is a special case of the 2^J design wherein all alternatives are part of each choice set presented to the respondent for evaluation. The procedure can also be easily extended to deal with Johnson's (1992) random-choice set designs.

As discussed in the introduction, choice-based conjoint models are typically fit at the aggregate level resulting in a single set of impact coefficients (or part-worths) being estimated for the entire sample of respondents. We propose to operationalize the notion of market segmentation by assuming the existence of $s = 1, \ldots, S$ homogeneous, latent segments of respondents with similar preference/part-worth structures. We denote the relative sizes of these segments via the S segment parameters $\alpha_s (0 \leq \alpha_s \leq 1, \Sigma_{r=1}^{S} \alpha_r = 1)$, which represent the proportion of respondents belonging to segment s.

Following Kamakura and Russell (1989), we express a respondent's choice probability in terms of the choice probabilities of these S segments. Conditional on a respondent belonging to segment s, the choice probability is given by

$$P_s(j\epsilon C_n) = \frac{\exp\left[\beta_{0js} + \sum_{k=1}^{K} X_{jk}\beta_{ks}\right]}{\sum_{a\epsilon C_n} \exp\left[\beta_{0as} + \sum_{k=1}^{K} X_{ak}\beta_{ks}\right]}, \tag{1}$$

where β_{0js} is the intrinsic utility of brand j to segment s and β_{ks} is the impact coefficient for attribute k in segment s. Hence, while the parameters of the choice-based conjoint model

are segment-specific, the composition of the segments are latent and unknown. The unconditional choice probability that a respondent chooses alternative j can be computed as

$$P(j \epsilon C_n) = \sum_{s=1}^{S} \alpha_s P_s(j \epsilon C_n), \tag{2}$$

where α_s, the size of segment s, may be construed as the *a priori* (or initial) probability of finding a respondent in segment s. Hence, our modeling framework entails a latent mixture of conditional logit models (cf. Kamakura and Russell, 1989; Ramaswamy and DeSarbo, 1990; Gupta and Chintagunta, 1994) to estimate the latent segments using the observed conjoint choice data. In particular, we modify the Kamakura and Russell (1989) scanner data response methodology to the choice-based conjoint arena where there are within-subject replications over (incomplete) choice set presentations. The Kamakura and Russell (1989) approach may be viewed as a special case of the proposed method when the size of the choice set (number of alternatives) is fixed over the consecutive choices for a given individual.

2.1. Model estimation

Given a sample of I respondents, the likelihood of the observed conjoint choice data can be formulated as

$$L = \prod_{i=1}^{I} \sum_{s=1}^{S} \alpha_s \prod_{n=1}^{N} \prod_{j \epsilon C_n} \left[\frac{\exp \left(\beta_{0js} + \sum_{k=1}^{K} X_{jk} \beta_{ks} \right)}{\sum_{a \epsilon C_n} \exp \left(\beta_{0as} + \sum_{k=1}^{K} X_{ak} \beta_{ks} \right)} \right]^{Y_{ijn}}, \tag{3}$$

where Y_{ijn} reflects the observed choice of respondent i for brand j in choice set n ($Y_{ijn} = 1$ if a choice is observed; 0 otherwise). The goal of the estimation is to maximize this likelihood, or equivalently, the *ln* likelihood) with respect to the segment-specific parameters $\mathbf{B} = ((\beta_{0js}, \beta_{ks}))$ and the S segment proportions $\mathbf{A} = (\alpha_s)$, subject to the constraint that $\Sigma_{r=1}^{S} \alpha_r = 1$ (see Wedel and DeSarbo, 1994, for a general discussion of latent-mixture regression models and their estimation).

On obtaining these parameter estimates, we can compute the posterior probability of membership, $\hat{R}_{is}$, for each respondent i (using Bayes rule) as

$$\hat{R}_{is} = \frac{\hat{\alpha}_s \prod_{n=1}^{N} \prod_{j \epsilon C_n} [\hat{P}_s(j)]^{Y_{ijn}}}{\sum_{s=1}^{S} \hat{\alpha}_s \prod_{n=1}^{N} \prod_{j \epsilon C_n} [\hat{P}_s(j)]^{Y_{ijn}}}, \tag{4}$$

where $\hat{P}_s(j)$ is the estimated choice probability for alternative j conditional on segment s, obtained by substituting the respective parameter estimates in expression (1). This feature accommodates situations where the consumers of a particular market segment display different preference patterns by allowing them to be fractional members of multiple segments. One could form discrete market segments, if desired, by simply assigning repondent i to the market segment whose value of $\hat{R}_{is}$ is highest. Thus, the proposed approach *simultaneously* estimates the conjoint part-worths and market segment membership with commonly collected choice-based conjoint data, without any additional data requirements.

Because the number of segments is unknown before analysis, we estimate the proposed model for varying values of S and use the consistent Akaike information criterion or CAIC (Bozdogan, 1987) to select a particular solution. The solution that minimizes the value of CAIC is chosen after ensuring its managerial interpretability. In addition, an entropy-based measure, E_s, is utilized to examine whether the derived segments are well separated (see Ramaswamy, DeSarbo, Reibstein, and Robinson, 1993, for a discussion of this measure).

3. Empirical application

To illustrate the proposed method, consider a commercial application of choice-based conjoint analysis conducted for a major packaged goods company as part of a new concept test. The product category entails frequently purchased food items offered in regular and diet forms except for local brands, a majority of which tend to be offered only in the regular form. For reasons of confidentiality, we cannot reveal the specific name of the product category or the identities of the brands. The product category is dominated by two national manufacturers, Alpha and Beta. Manufacturer Alpha has two separate product lines with distinct brand identities (denoted as Alpha 1 and Alpha 2). Under each of these two brand names, Alpha offered regular and diet versions of the product (denoted as R and D respectively). Hence, the four items offered by Alpha are denoted as Alpha 1-R, Alpha 1-D, Alpha 2-R, and Alpha 2-D. Manufacturer Beta, the sponsor of the study, has a single brand (Beta 1) in this product category that was offered in both the regular and diet forms. These are denoted as Beta 1-R and Beta 1-D respectively. Manufacturer Beta was considering introducing a new version of the diet product that capitalized on a key process technology resulting in a healthier alternative. Management wished to introduce this new product under a separate brand name, denoted as Beta N-D. Compared to Alpha, Beta placed relatively more emphasis on the corporate name in its packaging and its advertising, and consumers apparently had stronger associations between the manufacturer name Beta and its branded products.

To understand consumer tradeoffs between brand and manufacturer name, national and local brand, regular and diet product form, and price, a conjoint-choice task was designed using the existing six national brands (Alpha 1-R, Alpha 1-D, Alpha 2-R, Alpha 2-D, Beta 1-R, Beta 2-D), a local brand alternative (denoted as Local-R), and the new product concept (Beta N-D). These products were all of a standard package size. There were four price levels tested for each of the two product forms along with a fixed (low) price for the private label. The average price for the diet product form was \$1.44 compared to \$1.34 for the regular product form, reflecting the typical premium pricing for the former. The local brand alternative had a single fixed price of \$1.19. In other choice-based conjoint applications,

where the competitive set is not quite exhaustive or specific to a product category, this fixed alternative might represent a "no-choice or other" option. As in all choice-based conjoint models, there must be at least one constant alternative across choice sets to serve as the base for estimation purposes because the parameters of the underlying logit choice model are uniquely determined only up to an additive constant.

Two configurations (incomplete fractional factorials that together formed an orthogonal array) of sixteen choice sets were generated such that every choice set contained the eight brands, each at a specific price level. Hence, the choice-set design was fixed over a respondent's choice task. Each of the two incomplete fractional factorials was used with a separate group of respondents who were screened to ensure that they were category users and qualified for the study. The method of data collection was via personal interviews in several shopping malls. Respondents were asked to imagine that they were on a shopping trip in their usual store with intentions to make a purchase in the product category. Each respondent was then shown the sixteen choice sets, each one mimicking a set of brands on the store shelf. For each choice set, the respondent was asked to pick one brand at the price shown. In all, the conjoint choice data entail a sample of 600 respondents making a total of 9,600 choices. The mean price and choice shares for the brands are given in Table 1.

The common practice is to conduct an aggregate analysis of the conjoint-choice data using a conditional logit model. In such an analysis, all consumers are assumed to have the same preference structure, and only a single set of "part-worths" is estimated for the entire sample. This is equivalent to setting $S = 1$ in the method discussed in the previous section. Alternatively, we can conduct segment-level analyses using the proposed method by varying the value of S ($S > 1$). The CAIC heuristic was smallest for the four-segment solution. The value of the entropy-based measure, E_S, was 0.941 for the four-segment solution. This suggests that the posterior probabilities of membership were very well separated, and the market may be partitioned into four almost discrete segments (see Ramaswamy et al., 1993). The aggregate-level and segment-level estimates of the intrinsic brand utilities and price sensitivities, as well as the segment proportions for the four-segment solution, are given in Table 2. The log likelihood values for the aggregate analysis and the four-segment solution were $-17,391.7$ (CAIC = 34,864.8) and $-14,586.3$ (CAIC = 29,528.5), respectively.

Three of the four derived segments are relatively large, comprising 24.2 percent, 34.5 percent, and 34.5 percent of the sample. A key insight from Table 2 is that there is a small segment (6.8 percent) of respondents for whom the new concept is relatively unappealing. Further, this segment also does not have an intrinsic preference for national brands. Respon-

Table 1. Mean price and choice share

Brand	Mean Price	Choice Share
Alpha 1-R (Regular)	$1.34	8.1%
Alpha 1-D (Diet)	$1.44	9.6
Beta 1-R (Regular)	$1.34	7.7
Beta 1-D (Diet)	$1.44	13.7
Alpha 2-R (Regular)	$1.34	7.3
Alpha 2-D (Diet)	$1.44	11.9
Local-R (Regular)	$1.19	6.2
Beta N-D (New Diet Concept)	$1.44	35.6

Table 2. Choice-based conjoint estimates: Aggregate versus post hoc segmentation

		Segment-Level Analysis			
	Aggregate	Beta Segment (24.2%)	Diet Segment (34.5%)	Regular Segment (34.5%)	Local Segment (6.8%)
Intrinsic brand utilities (β_{0j}):					
Beta N-D	2.129	3.791	3.468	1.733	-0.936
Beta 1-D	1.150	3.026	2.264	0.621	-2.939
Alpha 2-D	1.003	-0.191^{ns}	2.991	0.794	-2.946
Alpha 1-D	0.779	0.546	2.582	0.752	-2.196
Alpha 1-R	0.476	-1.640	-1.093	1.692	-2.063
Beta 1-R	0.426	1.398	-0.120^{ns}	1.417	-3.620
Alpha 2-R	0.371	-1.449	-0.212	1.529	-2.145
Local-R	0.000	0.000	0.000	0.000	0.000
Price sensitivity (β_1)	-1.322	-0.630	-1.334	-1.768	-2.831
Segment size (α_s)		0.242	0.345	0.345	0.068

Note: All estimates are statistically significant at the 0.01 level, except for those denoted as ns.

dents in this small segment apparently have a high intrinsic utility for local brands and are extremely price sensitive. We label this small segment as a "local segment." In contrast, respondents in the third segment (34.5 percent) appear to have strong preferences for regular national brands, and are less price sensitive. This segment is therefore labeled as a "regular segment."

The new concept (Beta N-D) appears to have relatively less appeal in the regular and local segments. This is consistent with the fact that new concept is a diet product and is less likely to appeal to regular product buyers. In contrast, Beta N-D appeals more to respondents in the first two segments. These segments however exhibit interesting differences. The first segment (24.2 percent) appears to consist of respondents who have a high intrinsic preference for brands Beta N-D, Beta 1-D, and to some extent, Beta 1-R. This segment appears to be loyal to Beta, the manufacturer of the new concept. We label this segment as a "Beta segment." The second segment (34.5 percent) appears to consist of respondents who exhibit a high intrinsic preference for brands Beta N-D, Beta 1-D, Alpha 1-D, and Alpha 2-D, all of which have a diet product form. This segment is therefore labeled as a "diet segment." Furthermore, the Beta and diet segments exhibit lower price sensitivities than the regular and local segments. This suggests that consumers in the diet segment are willing to pay a premium for the benefits offered by the diet product form (lower calories and fat content), and that consumers in the Beta segment exhibit some loyalty to Manufacturer Beta, possibly due to the perceived higher quality communicated via its corporate advertising and product packaging.

Despite the presence of a potential new concept bias, we can make a relative comparison between the aggregate and segment-level results. It is apparent that the new concept appeals to two groups of consumers: those that are loyal to Beta (the Beta segment) and those that have a strong preference for diet brands (the diet segment). Further, there appear to be two groups of buyers of the regular product form: those that have a relatively strong pref-

erence for national brands (regular segment) and those that prefer local brand alternatives (local segment). The new diet concept does not seem to appeal to buyers of the regular product form. Finally, the Beta segment appears to be the least price sensitive. This segment's buyers seem to recognize the quality image and corporate identity created by Manufacturer Beta through its advertising and packaging. A summary of the market segmentation can be obtained by using expression (1) and mean brand prices to transform the within-segment parameter estimates in Table 2 into mean conditional brand choice probabilities for each segment. These are shown in Table 3. (Within-segment choice probabilities larger than aggregate choice probabilities are boxed.)

Besides the new product concept, it is evident that brands of the diet form (Beta 1-D, Alpha 2-D, Alpha 1-D) have relatively large shares in the diet segment, with Beta 1-D exhibiting a stronger presence within the Beta segment. Further, brands of the regular form (Alpha 1-R, Beta 1-R, alpha 2-R) have relatively large shares in the regular segment, while the local brand has a dominant position in the local segment. It is also evident from the above table that the new concept has a much larger share in the Beta and diet segments. Overall, the market seems to be structured according to a combination of brand and manufacturer name, product form (regular versus diet), and type of brand (national versus local).

In the present study, some ancillary data such as respondents' evaluation of the new concept, past purchase behavior (category and brand specific questions), attitudes toward health, fitness, diet, and nutrition, and background characteristics such as demographics were collected from respondents as part of the concept testing process. In general, by profiling the derived segments using such descriptive data, further insights into their composition can be obtained. In situations where the segment memberships exhibit some fuzziness, the posterior membership probabilities can be directly analyzed as a function of relevant descriptor variables (see Ramaswamy, DeSarbo, Reibstein, and Robinson, 1993, for an illustrative analysis of fuzzy posterior probabilities, albeit in a different modeling context). In the present context, given the high separation of the derived segments, discrete partitions of the sample were formed by assigning each respondent to the segment with the highest posterior

Table 3. Brand choice probabilities for each segment

	Beta Segment (24.2%)	Diet Segment (34.5%)	Regular Segment (34.5%)	Local Segment (6.8%)
Beta N-D	0.60	0.41	0.19	0.13
Beta 1-D	0.28	0.14	0.06	0.02
Alpha 2-D	0.01	0.25	0.08	0.02
Alpha 1-D	0.02	0.17	0.07	0.02
Alpha 1-R	0.00	0.00	0.21	0.06
Beta 1-R	0.06	0.01	0.16	0.01
Alpha 2-R	0.00	0.01	0.18	0.05
Local-R	0.02	0.02	0.04	0.69

probability. The segment proportions after modal allocations of the respondents are 0.242, 0.348, 0.343, and 0.067, which are almost identical to the estimated segment sizes. These discrete segments were then profiled using the ancillary data.

As might be expected, the Beta and diet segments exhibited higher top-two box scores for purchase intent and uniqueness of the new concept. Further, the Beta segment had a high incidence of purchase of Beta's products, which is again consistent with this segment's high intrinsic utility for these products. In contrast, the regular and local segments were indeed less favorable toward the new product concept, with the local segment having a very low incidence of purchase of Beta's products. Respondents belonging to the Beta and diet segments also appeared to have a strong orientation toward health and fitness and a basic commitment to watching their intake of calories and fat. This was in direct contrast to respondents in the regular and local segments, who appeared to be less concerned about such issues, with the latter segment having even lower concern. However, there was only a weak distinction between the segments with respect to demographics. The regular segment had relatively lower income respondents from somewhat larger households, while the local segment appeared to consist of "empty nesters" with larger-than-average incomes.

Overall, the segments had distinct evaluations of the new concept, attitudes toward fitness and nutrition, preferences for the new concept, and past category and brand purchase behavior that appeared to be consistent with the derived segmentation scheme, thereby lending some face validity to the results. Given the nature of the product, it was not surprising to find only weak differences in demographics. In general, background characteristics such as demographics have been found to be weakly associated with attribute preferences (Gupta and Chintagunta, 1994; Moore, 1980), and other background data such as perceived importance and appropriateness-in-use of various usage situations may provide better insights for food-related products. However, *a priori*, it may not be obvious *which set of background variables* should form the basis of the segments to enable disaggregate estimation of choice-based conjoint models. While it may be difficult to use background variables to form segments *a priori*, another option is to cluster respondents based on the observed-choice frequencies.

3.1. A priori segmentation based on observed-choice frequencies

We used a K-means cluster analysis of the observed-choice frequencies to form four segments. After splitting the sample into these four *a priori* segments, we estimated a separate choice-based conjoint model for each segment. The parameter estimates for each *a priori* segment (as well as the aggregate-level estimates) are shown in Table 4. There are three key observations from Table 4. First, the price sensitivity does not vary across the four *a priori* segments. Second, within each segment, the diet brands have larger intrinsic brand utilities than the regular brands; this ordering is similar to the aggregate case. Third, the four segment sizes from the K-means cluster analysis are about equal. Overall, the *a priori* segment-level results do not exhibit any striking variation across the segments and are similar to the aggregate-level estimates. What about the choice shares within each segment? While one would expect variation in the observed within-segment shares (given that the cluster analysis was based on the observed individual choice frequencies), we found hardly

Table 4. Choice-based conjoint estimates: Aggregate versus *a priori* segmentation

		Segment-Level Analysis			
	Aggregate	Segment 1 (21.2%)	Segment 2 (25.8%)	Segment 3 (25.7%)	Segment 4 (27.3%)
Intrinsic brand utilities (β_{0j}):					
Beta N-D	2.129	1.760	1.874	2.035	1.718
Beta 1-D	1.150	1.056	1.102	1.290	1.093
Alpha 2-D	1.003	0.346	0.433	0.613	0.366
Alpha 1-D	0.779	0.360	0.465	0.649	0.410
Alpha 1-R	0.476	−0.453	0.035	0.227	−0.058
Beta 1-R	0.426	−0.254	−0.042	0.158	−0.097
Alpha 2-R	0.371	−0.138	0.206	0.395	0.091
Local-R	0.000	0.000	0.000	0.000	0.000
Price sensitivity (β_1)	−1.322	−1.090	−1.144	−1.156	−1.123

any variation in the estimated segment-level conditional choice probabilities. Since the cluster analysis and the within-segment choice-based conjoint analysis *do not maximize the same objective function* in the *a priori* case, there is no assurance that the *a priori* segments will exhibit different intrinsic brand utilities and price sensitivities. Clustering the choice frequencies results in *a priori* segments with less variation in brand choices and increases the chances of obtaining insignificant parameter esitmates. In fact, the log likelihood based on the *a priori* segmentation results was −17,474.7, which is no better than the aggregate case. The congruence between the *a priori* classification of respondents and the *post hoc* discrete segment allocation (after appropriate permutation) is only 57 percent. Further, as noted by DeSarbo, Wedel, Vriens, and Ramaswamy (1992), different clustering procedures can potentially lead to different *a priori* segmentation schemes and consequently, different segment-level results.

   The fundamental goal of market segmentation with conjoint analysis is to capture market heterogeneity in brand and attribute preferences and price sensitivity. In traditional ratings-based conjoint analysis, this is afforded through the potential of conducting individual-level estimation and then clustering the resultant individual-level part-worths to form segments. In a choice-based conjoint setting, it is inefficient to estimate individual-level parameters with a limited number of choices. The proposed approach is an attractive option for carrying out market segmentation with choice-based conjoint analysis.

4. Conclusion

We have discussed an approch for market segmentation with choice-based conjoint analysis. The inability to derive market segments has heretofore been a primary disadvantage of choice-based conjoint analysis (Struhl, 1994). We overcome this disadvantage by accommodating consumer heterogeneity via multiple segments of consumers possessing different utility structures. A major benefit of the proposed method is that the segments are derived empirically using commonly collected conjoint-choice data without recourse to any *a priori* specification of the nature of the heterogeneity (typically based on ancillary background

data). The number of segments, their sizes and utility structures are all derived empirically, in a maximum-likelihood framework.

We demonstrated the proposed method using commercial conjoint-choice data gathered in a new concept test. Our approach, however, is not restricted to this type of application and can be used in other settings involving different designs. In general, our approach is applicable in any instance where conjoint-choice data have been gathered, although, as discussed, researchers may be currently analyzing the data in the aggregate or at the level of *a priori* segments. Clearly, further testing of the approach is necessary to establish its usefulness to practitioners. Since no validation profiles were collected in the commercial application, we could not perform holdout analyses to compare share predictions based on our approach versus an aggregate-level approach. There is some empirical evidence from the choice modeling literature that models which capture heterogeneity, particularly via multiple market segments, perform better than aggregate-level models in terms of predicting both aggregate and individual-level choice shares (Ramaswamy and DeSarbo, 1990). It would be fruitful to perform a cross-validation assessment of our approach and other alternatives to market segmentation (different *a priori* specifications of heterogeneity), similar in spirit to Green and Helsen (1989) who compare alternatives to individual-level ratings-based conjoint models.

References

Bozdogan, H. (1987). "Model Selection and Akaike's Information Criterion: The General Theory and its Analytical Extensions." *Psychometrika*, 52, 345–370.

DeSarbo, W.S., M. Wedel, M. Vriens, and V. Ramaswamy. (1992). "Latent Class Metric Conjoint Analysis." *Marketing Letters*, 1992, 273–288.

Elrod, T., J.J. Louviere, and K.S. Davey. (1992). "An Empirical Comparison of Ratings-Based and Choice-Based Conjoint Models." *Journal of Marketing Research*, 29 (August), 368–377.

Green, P.E., and A.M. Krieger. (1991). "Segmenting Markets with Conjoint Analysis." *Journal of Marketing*, 55 (October), 20–31.

Green, P.E., and K. Helsen. (1989). "Cross-Validation Assessment of Alternatives to Individual-Level Conjoint Analysis: A Case Study." *Journal of Marketing Research*, 26, 346–350.

Gupta, S., and P.K. Chintagunta. (1994). "On Using Demographic Variables to Determine Segment Membership in Logit Mixture Models." *Journal of Marketing Research*, 31, 128–136.

Johnson, R. (1992). *Sawtooth Software: The CBC System*. Evanston, IL: Sawtooth.

Kamakura, W.A., and G.J. Russell. (1989). "A Probabilistic Choice Model for Market Segmentation and Elasticity Structure." *Journal of Marketing Research*, 26, 379–390.

Louviere, J.J., and G.G. Woodworth. (1983). "Design and Analysis of Simulated Choice or Allocation Experiments: An Approach Based on Aggregate Data." *Journal of Marketing Research*, 20, 350–367.

Moore, W.M. (1980). "Levels of Aggregation in Conjoint Analysis: An Empirical Comparison." *Journal of Marketing Research*, 17, 516–523.

Ramaswamy, V., and W.S. DeSarbo. (1990). "SCULPTRE: A New Methodology for Deriving and Analyzing Hierarchical Product-Market Structures from Panel Data." *Journal of Marketing Research*, 27 (November), 418–427.

Ramaswamy, V., W.S. DeSarbo, D.J. Reibstein, and W.T. Robinson. (1993). "An Empirical Pooling Approach for Estimating Marketing Mix Elasticities with PIMS Data." *Marketing Science*, 12(1) (Winter), 103–124.

Struhl, S. (1994). "Discrete Choice Modeling: Understanding a 'Better Conjoint than Conjoint.' " *Quirk's Marketing Research Review*, 8, 6, 12+.

Wedel, M. and W.S. DeSarbo. (1994). "A Review of Recent Developments in Latent Class Regression Models," in *Advanced Methods of Marketing Research*, R.P. Bagozzi (ed.), 352–388. Cambridge, MA: Blackwell.

Wittink, D.R., and P. Cattin. (1989). "Commercial Use of Conjoint Analysis: An Update." *Journal of Marketing*, 53, 91–96.