Capturing Heterogeneity in Customer Satisfaction Models:

A Finite Mixture PLS Approach[1]

By

Carsten H. Hahn, Michael D. Johnson and Andreas Herrmann[2]

August, 2000

Capturing Heterogeneity in Customer Satisfaction Models:

A Finite Mixture PLS Approach

Abstract

A new approach for capturing unobserved heterogeneity in customer satisfaction models is proposed based on partial least squares (PLS). The method uses a modified finite-mixture distribution approach. An empirical analysis using quality, customer satisfaction and loyalty data for convenience stores illustrates the advantages of the new method vis-à-vis a traditional market segmentation scheme based on well known grouping variables. The results confirm the assumption of heterogeneity in the individuals' perception of the antecedents and consequences of satisfaction and their relationships. The results also illustrate how the finite-mixture approach complements and provides insights over and above a traditional segmentation scheme.

Introduction

The successful allocation of marketing resources requires that managers understand just what factors drive customer satisfaction and subsequent loyalty. Customer satisfaction is a major outcome of marketing activity and serves to link customers' purchase and consumption experiences with post-purchase phenomena such as attitude change, repeat purchase, and brand loyalty (Oliver 1997). Managers in the service industry are placing particular emphasis on customer satisfaction to enhance customer loyalty for long-term profitability (Heskett et al. 1997). Consequently, a marketing manager needs to know just what quality or price factors to emphasize or invest in to improve satisfaction.

At the same time, any analysis of the drivers of satisfaction must build upon, or take into account, segment-level differences. For example, satisfaction and loyalty to a convenience store may be driven primarily by "convenience" for a frequent user and "safety" for an infrequent

- 2 -

user. One traditional approach is to use separate marketing research (interviews, focus groups and surveys) to identify a priori segments on which the analysis of satisfaction data is built. A more recent trend in marketing research is to determine the segments when analyzing customer data using a latent class or finite mixture approach (Jedidi, Jagpal and DeSarbo 1997a; Ramaswamy et al. 1993).

There are two popular causal-modeling methods for estimating satisfaction models, covariance structure analysis using programs such as LISREL and partial least squares (PLS; Gustafsson and Johnson 1997). While the finite mixture approach has been added and applied to covariance structure analysis (Jedidi, Jagpal and DeSarbo 1997a, 1997b), it has not been integrated with PLS. This is despite the fact that, when operationalizing a quality and satisfaction model, PLS has distinct advantages over other methods. The goal of this research is to integrate the advantages of PLS with the advantages of a finite mixture approach to market segmentation. The integration is unique in that it leverages the advantages of a least-squared procedure when operationalizing a satisfaction model and the advantages of a maximum likelihood-based approach when deriving market segments.

We compare and contrast the approach with a more traditional a priori segmentation scheme using data from a national convenience store survey. The new approach both complements that traditional segmentation scheme and provides unique insights into the drivers of customer satisfaction. We begin by describing customer satisfaction modeling and approaches to estimating satisfaction models. We then describe our Finite Mixture PLS approach, our empirical study and results. The advantages and disadvantages of the approach vis-à-vis traditional market segmentation are then described and discussed.

Customer Satisfaction Modeling

Customer satisfaction models typically include the concrete attributes that describe a product or service, the benefits or consequences these attributes provide customers, a customer's overall evaluation of their purchase and consumption experience (customer satisfaction), and the behavioral intentions or behaviors that result (such as repurchase, product recommendations or word-of-mouth, cross-selling, or price tolerance). These models rest heavily on expectancy-value model formulations (Bagozzi 1982, 1994), where beliefs about the consumption experience (quality dimensions and price) affect customer satisfaction as a type of overall evaluation or attitude, which in turn affects customers' behavioral intentions and behaviors.

A key feature of satisfaction models is that the benefit, satisfaction, and loyalty constructs in the models are inherently abstract or latent variables. The most common way to empirically measure these latent variables is through the use of multiple concrete proxies or measurement variables. Benefits are measured using their attributes, satisfaction is measured using different overall evaluation standards (such as overall satisfaction, overall performance versus expectations, overall performance versus an ideal), and loyalty is measured using behavioral intentions (such as the likelihood of repurchase or recommendation to others).

Statistical estimation of a satisfaction model must accommodate the fact that the model is a network of cause-and-effect relationships (as from quality, to satisfaction, to loyalty) that contains latent variables. There are two popular methods for estimating models of this type, partial least squares (PLS) and covariance structure analysis (CSA; Fornell 1987). The methods are more complementary than competing; their use should depend on the purpose of the analysis. Covariance structure analysis is particularly well suited to testing among multiple causal models based on strong theory and data. When, however, the goal is to operationalize an existing model,

such as a company's customer satisfaction model, PLS is preferred (Gustafsson and Johnson 1997; Steenkamp and van Trijp 1996).

PLS is essentially an iterative estimation procedure that integrates principal-components analysis with multiple regression (Fornell and Cha 1994). Whereas the objective of CSA is to explain covariance, the objective of PLS is to explain variance in the endogenous variables in a satisfaction model that have bottom-line managerial relevance (satisfaction, loyalty, profit). Unlike CSA, the latent variables in PLS are easily operationalized as principal components or weighted indices of the measurement variables. This provides managers with explicit benchmarks for evaluating their performance. When this performance information is combined with the impact scores from the regression estimates, managers have both the impact and performance information that they need to make key resource allocation decisions (Martilla and James 1977).

There are other reasons to use PLS in a satisfaction-modeling context. Bagozzi and Yi (1994) argue that PLS is preferred over CSA (LISREL) when: (1) sample sizes are small, (2) the data to be analyzed is not multivariate normal, as when distributions are highly skewed, and (3) improper or non-convergent results are likely, as when estimating a complex model with many variables and parameters. Satisfaction models often use small samples, especially at the segment level. Quality and customer satisfaction data is also marked by large negative skewness (Fornell 1995). Furthermore, satisfaction models are often large and complex, involving multiple abstract benefits and dozens of attributes. Thus PLS is arguably superior to CSA when *operationalizing* a satisfaction model. One trade-off is that PLS tends to underestimate path coefficients and overestimate loadings (Bagozzi and Yi 1994). As Bagozzi and Yi argue, however, this means that the significant results of a PLS analysis can be given more credence because the test is more

- 5 -

conservative. Other limitations of PLS are that jackknife or bootstrap procedures are needed to obtain estimates for the standard errors of the parameter estimates and, because PLS is a limited-information estimation method, its estimates are not as efficient as full-information estimates (Bagozzi and Yi 1994; Fornell and Bookstein 1982). Overall, however, there are clear reasons for integrating the advantage of PLS with the advantages of the finite mixture approach in a satisfaction context.

Capturing Heterogeneity in Satisfaction Models

When estimating structural equation models, researchers frequently treat data as if it were collected from a single population (Muthén 1989). This is unlikely to be the norm in customer satisfaction research. In multidimensional expectancy value models, customers from different market segments can have very different belief structures (Bagozzi 1982). Thus the impact that different drivers have on satisfaction, and their level of performance, likely varies from segment to segment. Consider a simple model in which two drivers, price (P) and quality (Q), impact satisfaction (S). The constructs are measured indirectly using multiple indicators, but our focus is on the two path coefficients in the inner model, where price impacts satisfaction (P → S) and quality impacts satisfaction (Q → S). In an aggregate model, say the estimated impacts are both equal to 0.3 and the research concludes that quality and price are equally important. However, suppose there are two unobserved segments of equal size. The individuals belonging to each segment are homogeneous in how quality and price impact satisfaction. In segment 1 quality has five times the impact of price (0.5 versus 0.1), while in segment 2 the opposite is true (0.1 versus 0.5). In reality, the population is composed of a quality-oriented segment and a price-oriented segment. The aggregate structural equation model leads to meaningless results where managers derive the wrong interpretation. As a result, marketing and quality resources could be poorly

allocated and achieve little. In this case, the result would be a "value" positioning that appeals to neither the quality nor price-oriented segments.

Typically, heterogeneity in structural equation models has been addressed by assuming that consumers can be assigned to segments a priori of the basis of demographic variables, usage levels, or other proxies for the underlying segments (Wedel and Kamakura 1999). A limitation of the a priori approach is that heterogeneity is often not captured adequately by well-known observable variables. Jedidi, Jagpal and DeSarbo (1997a, 1997b) propose a new approach based on CSA where heterogeneous groups are identified simultaneously with the structural equation model using a finite mixture framework. Arminger and Stein (1997) propose a more general method based on covariance structure estimates.

An alternative is to develop a hierarchical Bayesian methodology for treating heterogeneity in structural equation models (Ansari, Jedidi and Jagpal 1999). An important advantage of this methodology is that it automatically provides individual-specific estimates of model parameters and factor scores. This is an interesting source for marketing managers who want to implement a relationship-marketing concept based on individual customer-to-supplier relationships. However, the researcher requires some meaningful, a priori information about the parameters and more than one observation from at least some individuals.

Again, our goal is to combine the advantages of PLS with the finite mixture approach. Figure 1 provides a taxonomy of methods that seek to capture heterogeneity in structural equation models and shows where our proposed approach fits in the taxonomy. Substantively, the approach allows the marketing manager to perform response-based market segmentation where all consumers or customers in a segment are homogeneous in terms of the model's path coefficients. As we will show, the approach complements a priori segmentation by capturing

heterogeneity within existing, well-known segments. Methodologically, the approach contributes to marketing research by allowing researchers to detect unobservable, discrete moderating factors that account for heterogeneity among consumers based on the prediction-oriented PLS approach. This approach combines the advantages of predicting path coefficients, using PLS, with the maximum likelihood estimation of a finite mixture model. Conceptually the approach expands the a priori segmentation methods to prediction-oriented structural equation models.

- insert Figure 1 about here -

The Finite Mixture Partial Least Squares Approach

Wold (1966) originally developed the PLS approach as an algorithm for least squares (LS) estimation of path models with latent variables. Each latent variable (LV) is indirectly observed by a block of manifest variables (MVs). PLS predicts the linear conditional expectation relationship between dependent and independent variables. As this approach is based on predictor specification, PLS is quite different from covariance structure analysis like LISREL that focuses on a causality concept based on accounting for covariances.

A path model with latent variables (structural equation model) consists of an inner model (inner relations, structural model, substantive part) and an outer model The inner model depict the relationships among the latent variables as posited by substantive theory.

Let

I = the subject (observation, individual) i with $i = 1,..., N$

η_i = the vector of the endogenous variables in the inner model for subject i

ξ_i = the vector of the exogenous variables in the inner model for subject i

The inner relations can be expressed by:

$$B\eta_i + \Gamma\xi_i = \zeta_i \qquad (1)$$

where $B(Q \times Q)$ and $\Gamma(Q \times P)$ are path coefficient matrices with Q = number of endogenous variables, P = number of exogenous variables and ζ_i' is a random vector of residuals.[3]

Outer relations define the relationships between the manifest variables (indicators) and the latent variables (components). Two kinds of outer relationships can be specified: reflective and formative. PLS allows for either type of relationship.

Let

x_i = the vector of observed measures for the exogenous LV for subject i

y_i = the vector of observed measures for the endogenous LV for subject i

The outer relations for the reflective (outward) model can be expressed by:

$$y = \Lambda_y \eta + \varepsilon_y \qquad (2a)$$

$$x = \Lambda_x \xi + \varepsilon_x \qquad (2b)$$

where $\Lambda_y(K \times Q)$ and $\Lambda_x(P \times L)$ are the matrices of loadings that relate the latent variables to their measures where K is the number of indicators for endogenous variables, L is the number of indicators for exogenous variables and the ε's are the residuals and usually interpreted as measurement errors or noise.

In the formative case, the relationships between the LVs and their indicators are defined as:

$$\eta = \pi_\eta + \delta_\eta \qquad (3a)$$

$$\xi = \pi_\xi + \delta_\xi \qquad (3b)$$

where the π's are the multiple regression coefficients and δ's are the residuals from regressions.

---

[3] The more familiar formulation of the inner relations is: $\eta = B^*\eta + \Gamma^*\xi + \zeta$. Equation (1) can be easily obtained by setting $B = (I - B^*)$ and $\Gamma = -\Gamma^*$ and hence the two formulations are equivalent.

The usual PLS algorithm predicts B, Γ, the Λ's and the π's with an iterative scheme of partial least squares and calculates the scores of η and ξ for every individual. The result is an aggregate predictor specification based on the constraints of B and Γ for the whole sample.

Conceptually, heterogeneity in a satisfaction model is concentrated in the path coefficients that relate quality factors and price to satisfaction and subsequent loyalty (Johnson and Gustafsson 2000). The proposed model is an approach to capture this heterogeneity. It assumes that η_i is distributed as a finite mixture of conditional multivariate normal densities, $f_{i|k}(\bullet)$:

$$\eta_i \sim \sum_{k=1}^{K} \rho_k f_{i|k}(\eta_i \mid \xi_i, B_k, \Gamma_k, \Psi_k) = \sum_{k=1}^{K} \rho_k \left[\frac{|B_k|}{(2\pi)^{Q/2} |\Psi_k|^{1/2}} \exp\left(-\frac{1}{2} (B_k\eta_i + \Gamma_k\xi_i)'\Psi_k^{-1}(B_k\eta_i + \Gamma_k\xi_i) \right) \right] \qquad (4)$$

where:

k	= 1,..., K latent classes;
m	= 1,..., Q number of endogenous variables;
j	= 1,..., P number of exogenous variables;
B_k	= $((\beta_{rmk}))$, the ($Q \times Q$) matrix of endogenous variables coefficients for latent class k ($r = 1,...,$ Q);
Γ_k	= $((\gamma_{mjk}))$, the ($Q \times P$) matrix of exogenous variables coefficients for latent class k;
Ψ_k	= the ($Q \times Q$) matrix with the variances for each regression of the inner model on the diagonal and zero else;
ρ	= $(\rho_1, ..., \rho_K)$, a vector of the K mixing proportions of the finite mixture (of which $K - 1$ are independent) such that $\rho_k > 0$ and $\sum_{k=1}^{K} \rho_k = 1$.

Suppose, the η_i vectors are independent; the likelihood function for the N vectors $(\eta_1,..., \eta_N)$ is given by:

$$L = \prod_{i=1}^{N}\left[\sum_{k=1}^{K}\rho_k\left[\frac{|\mathbf{B}_k|}{(2\pi)^{Q/2}\,|\Psi_k|^{1/2}}\quad \exp\left(-\frac{1}{2}\;(\mathbf{B}_k\eta_i+\Gamma_k\xi_i)'\Psi_k^{-1}(\mathbf{B}_k\eta_i+\Gamma_k\xi_i)\right)\right]\right] \tag{5}$$

The mixing proportions ρ can be construed as prior probabilities of any subject belonging to the K latent classes. The posterior probability of membership for subject i in class k $\left(\hat{P}_{ik}\right)$ can be computed using Bayes' theorem, conditional on the estimates of the class-specific parameters $\hat{\rho}_k, \hat{\mathbf{B}}_k, \hat{\Gamma}_k, \hat{\Psi}_k$ via:

$$\hat{P}_{ik} = \frac{\hat{\rho}_k f_{i|k}(\eta_i \mid \xi_i, \hat{\mathbf{B}}_k, \hat{\Gamma}_k, \hat{\Psi}_k)}{\sum_{k=1}^{K}\hat{\rho}_k f_{i|k}(\eta_i \mid \xi_i, \hat{\mathbf{B}}_k, \hat{\Gamma}_k, \hat{\Psi}_k)} \tag{6}$$

Identification of the Model

Mixtures of multivariate normal densities are typically identified (McLachlan and Basford 1988), but the model specified in Equation 4 is not identified if all elements in $\Gamma = (\Gamma_1, ...,\Gamma_K)$, $B = (\mathbf{B}_1, ..., \mathbf{B}_K)$ and $\Psi = (\Psi_1, ..., \Psi_K)$ are free. Identification in this context requires placing restrictions on model parameters. The most common restrictions set some elements of Γ, B and Ψ to zero or some other constant, whereas others entail the imposition of equality or inequality constraints on parameters (Jöreskog 1977).

Estimation of the Model via the EM-Algorithm

The likelihood of the model developed in the previous section can be maximized using the EM algorithm (Dempster, Laird and Rubin 1977; McLachlan and Krishnan 1997). The algorithm contains an expectation part (E-step) and a maximization part (M-step). It should be mentioned that another optimization routine such as Newton-Raphson or Fletcher-Reeves could

be used to maximize the likelihood-function. Convergence is not ensured with the two latter methods. The EM algorithm is attractive because it can be programmed easily and convergence is ensured. The estimation procedure can be described as a two-stage process. In the first stage a PLS solution is estimated based on the aggregate sample with the aim of obtaining predictor scores for the latent variables, η and ξ, for each respondent individually. In the second stage the predicted scores of the latent variables are used as dependent and independent variables for a set of regressions of the inner model, defined by the constraints of B and Γ respectively. Every endogenous variable reflects a regressant of an OLS regression, whereas the regressors come from a subset of endogenous and exogenous variables. All regression equations are computed independently according to the PLS assumption. Consequently the matrix for latent class k, Ψ_k, is a diagonal matrix with the variances of the partial regressions on the diagonal. Our segmentation approach relaxes the second stage by implementing a finite mixture model with this set of regression equations. The modification of the M-step is described later.

In order to present an EM formulation, we introduce nonobserved data via the indicator function:

z_{ik} = 1   if subject i belongs to class k,

        = 0   otherwise.

We assume that the nonobserved data in the vector $z_i = (z_{i1}, ..., z_{iK})$ are independently and identically multinomially distributed with probabilities ρ. The joint likelihood of η_i and z_i is

$$L_i(\eta_i, z_i; \xi_i, B_k, \Gamma_k, \Psi_k, \rho_k) = \prod_k [\rho_k f(\eta_i | \xi_i, B_k, \Gamma_k, \Psi_k)]^{z_{ik}} \tag{7}$$

The complete likelihood over all subjects is

$$L = \prod_i \prod_k [\rho_k f(\eta_i | \xi_i, B_k, \Gamma_k, \Psi_k)]^{z_{ik}} \tag{8}$$

and the log-likelihood is

$$\ln L = \sum_i \sum_k z_{ik} \ln(f(\eta_i | \xi_i, B_k, \Gamma_k, \Psi_k)) + \sum_i \sum_k z_{ik} \ln \rho_k. \tag{9}$$

The matrix $Z = (z_1, ..., z_I)$ is considered as missing data.

The E-Step

The EM-algorithm starts with an E-step, where the expectation of $\ln L$ is evaluated over the conditional distribution of the nonobserved data Z given the predicted values of η_i and ξ_i of the observed data y_i and x_i, and the provisional estimates (B^*, Γ^*, Ψ^*, and ρ^*) of the parameters B, Γ, Ψ, and ρ respectively. These estimates can be calculated from a random sample of membership probabilities of P_{ik} or can be set from the analyst based on assumptions and/or prior knowledge about the classes and the coefficients.

The expectation of the likelihood function is

$$E(\ln L; \xi, \rho = \rho^*, B = B^*, \Gamma = \Gamma^*, \Psi = \Psi^*)$$
$$= \sum_i \sum_k E(z_{ik}; \xi_i, \rho^*, B^*, \Gamma^*, \Psi^* | \eta_i) \ln(f(\eta_i | \xi_i, \rho^*_k, B^*_k, \Gamma^*_k, \Psi^*_k)) \tag{10}$$
$$+ \sum_i \sum_k E(\xi_i, \rho^*, B^*, \Gamma^*, \Psi^* | \eta_i) \ln \rho^*_k.$$

The conditional expectation of z_{ik} can be calculated as

$$E(\ln L; \xi, \rho = \rho^*, B = B^*, \Gamma = \Gamma^*, \Psi = \Psi^*)$$
$$= \rho^*_k f(\eta_i | \xi_i, B^*_k, \Gamma^*_k, \Psi^*_k) / \sum_k \rho^*_k f(\eta_i | \xi_i, B^*_k, \Gamma^*_k, \Psi^*_k). \tag{11}$$

Comparing (11) with (6) reveals that the posterior membership probability P_{ik}^* for subject i in class k evaluated with provisional estimates is

$$P_{ik}^* = E(z_{ik}; \xi_i, B^*, \Gamma^*, \Psi^* | \eta_i). \tag{12}$$

The nonobserved data in matrix Z are replaced by the posterior probabilities calculated on the base of provisional estimates. Thus equation (10) becomes

- 13 -

$$E(\ln L; \xi, \rho = \rho^*, B = B^*, \Gamma = \Gamma^*, \Psi = \Psi^*)$$

$$= \sum_i \sum_k P_{ik}^* \ln(f(\eta_i | \xi_i, \rho^*_k, B^*_k, \Gamma^*_k, \Psi^*_k)) \tag{13}$$

$$+ \sum_i \sum_k P_{ik}^* \ln \rho_k^*.$$

The M-step

In the M-step we maximize equation (9) with respect to the parameters subject to the restriction $\rho > 0$ *and* $\sum_k \rho_k = 1$, conditional on the new provisional estimates of z_{ik}, in order to obtain revised parameter estimates. These revised estimates are then used in the subsequent E-step to calculate new estimates of P_{ik}. These estimates are used as expectations of z_{ik} in the next M-step to get new estimates of the parameters and so forth.

In our approach the M-step contains a number of independent OLS regressions, one for each regression in the inner model. The regressions of the inner model reveal the relationships between the m endogenous variables (as dependent variables) and the exogenous and endogenous variables (as independent variables) of the model. The relationships are defined via B and Γ. Thus for each endogenous variable as a dependent variable an OLS regression is calculated in the M-step. We use the Maximum Likelihood Estimator of the coefficient and the variance, that is identical to the Least Squares Prediction in the OLS case.

Let

m = number of independent regressions in the inner model

A_m = number of exogenous variables as regressors in regression m

B_m = number of endogenous variables as regressors in regression m

Y_{mi} = the value of the regressant for regression m of individual i

X_{mi} = the value of the regressor $(A_m + B_m \times 1)$-vector for regression m of individual i

- 14 -

We obtain the parameters of the regression for endogenous variable *m* with

$$Y_{mi} = \eta_{mi}$$

$$X_{mi} = (E_{mi}, N_{mi})'$$

where

$$E_{mi} = \begin{cases} (\xi_1, ..., \xi_{A_m}) & \text{if } A_m >= 1, a_m = 1, ..., A_m \text{ and } \xi_{a_m} \text{ is regressor of regression } m \\ () & \text{else.} \end{cases}$$

$$N_{mi} = \begin{cases} (\eta_1, ..., \eta_{B_m}) & \text{if } B_m >= 1, b_m = 1, ..., B_m \text{ and } \eta_{b_m} \text{ is regressor of regression } m \\ () & \text{else.} \end{cases}$$

and the closed-form OLS analytic expressions for τ_{mk} and ω_{mk}

$$\tau_{mk} = \left[\sum_i P_{ik}(X'_{mi}X_{mi}) \right]^{-1} \left[\sum_i P_{ik}(X'_{mi}Y_{mi}) \right],$$

(14)

with

$$\tau_{mk} = \left(\left(\left(\beta_{a_m mk} \right) \right), \left(\left(\gamma_{b_m mk} \right) \right) \right)',$$

and

$$\omega_{mk} = \left[\sum_i P_{ik}(Y_{mi} - X_{mi}\tau_{mk})(Y_{mi} - X_{mi}\tau_{mk})' \Big/ I\rho_k \right],$$

(15)

with

$$\omega_{mk} = cell\ (m \times m)\ of\ \Psi_k.$$

The result of each independent regression serves as a new provisional estimate for the next E-step iteration of the EM-algorithm. The E-step and the M-step are successively applied until no further improvement in the ln-likelihood-function is possible based on a pre-specified convergence criterion. Hence, although convergence to at least a locally optimum solution is

guaranteed, different starting values of the parameters must be used to investigate the potential occurrence of the local optimum.

Model Selection

When applying the above model to data, the actual number of classes K is unknown and must be inferred from data. The problem of identifying the number of classes is still without a satisfactory statistical solution (Wedel and Kamakura 1999). The likelihood ratio test statistic for example is not valid in a mixture model, because it is not asymptotically distributed as chi-square. Bozdogan and Sclove (1984) propose using Akaike's (1974) Information Criteria (AIC) for determining the number of classes in a mixture model:

$$AIC_K = -2\ln L + cN_K, \tag{16}$$

with $c = 2$ is a constant and N_K is the number of free parameters:

$$N_K = (K-1) + KR + KQ. \tag{17}$$

R is the number of predictor variables in all regressions of the inner model. The constant c in AIC imposes a penalty on the likelihood, which weighs the increase in fit (more parameters yield a higher likelihood) against the additional number of parameters estimated (Wedel and Kamakura 1999).

Two criteria penalize the likelihood more heavily: Schwarz's (1978) Bayesian information criteria

$$BIC_K = -2\ln L + \ln IN_K, \tag{18}$$

where $c = \ln I$ and the consistent Akaike information criteria computed as

$$CAIC_K = -2\ln L + \ln(I+1)N_K, \tag{19}$$

where $c = (lnI + 1)$. However all measures discussed above are heuristics for model selection. To assess the separation of the segments, an entropy statistic (Ramaswamy et al. 1993) can be used to investigate the degree of separation in the estimated individual class probabilities.

$$EN_K = 1 - \left[\sum_i \sum_k - P_{ik} \ln P_{ik} \right] \bigg/ I \ln K.$$

(20)

EN_K is a relative measure and is bounded between 0 and 1. Values close to 1 indicate that the derived classes are well separated. In addition the entropy measure indicates whether a solution is interpretable or not. For example a solution with goods heuristics and a bad entropy measure, say $EN_K = 0$, can not be interpreted accurately. The segments are "fuzzy," which means that only 'parts' of the subjects belong to a class. The fuzziness of any derived class memberships makes the managerial implications equally "fuzzy."

Empirical Application

We illustrate the Finite Mixture PLS approach using a national survey of customers' perceived quality, satisfaction and loyalty with convenience stores. The survey was sponsored by the National Association of Convenience Stores (NACS) and based on a representative cross-section of convenience store customers and stores in the United States. The interview methodology (computer aided telephone interviews) and random sampling procedure was the same as that used for the American Customer Satisfaction Index (ACSI) survey (see Fornell et al. 1996). The data were collected in December of 1998, and the sample included 1,025 customers who were selected to be representative of the demographic profile of convenience store consumers. In terms of demographics, 42.4% were male and 57.6% were female. The age range was from 18-81 with a median age of 37. The sample was also broadly distributed across income and education levels.

The NACS satisfaction model is presented in Figure 2. The Figure shows nine latent variables, benefits or consequences that are immediate antecedents of satisfaction in the model. Satisfaction, in turn, affects customer loyalty, both of which are also latent variables in the model. Each latent variable is operationalized using multiple proxies or survey measures rated on 10-point scales (see Table 1). The satisfaction measures are the same as those used in the ACSI survey, while the loyalty measures are rated likelihoods of revisiting the convenience store and recommending it to others.

- insert Figure 2 and Table 1 about here -

Aggregate results

The aggregate PLS results are shown in Table 2. The values of the inner model calculated with PLS are equal to these of our new approach with $K = 1$. According to the aggregate path coefficients, the largest drivers of satisfaction are perceived safety (0.193), store layout (0.174), prices (0.168) and separate take out (0.152). The smallest drivers of satisfaction are products (0.039) and motorist services (0.039). This suggests that, on the whole, both products and gas services are relatively undifferentiated across convenience stores. Satisfaction has a large and significant impact on loyalty in the model (0.625).

- insert Table 2 about here -

The aggregate results for $K = 1$ provide important benchmark values for the goodness of fit measures AIC, BIC and CAIC, which are shown in the first row of Table 3. The next section discusses the constraining of parameters and results for K greater than one.

- insert Table 3 about here -

Disaggregate Finite Mixture Results

We applied our new approach for a varying number of classes K. The impacts should be greater or equal to zero (for both the aggregate and disaggregate models) as all of the survey questions are valenced in the same direction (higher values are more attractive, such as higher quality or more attractive prices). If we assume that all drivers of satisfaction are independent of each other, the impacts can be interpreted as the increase in satisfaction that results from an increase in any particular price or quality driver. Our initial disaggregate solutions contained some small, negative coefficients for the drivers. With the assumption of independence, an interpretation would be difficult. For example, a negative impact of service to satisfaction would mean that higher quality service lowers overall satisfaction. To prevent such non-interpretable solutions, we constrained our coefficients to be equal or higher than zero. Consequently we only obtain local but interpretable optima for our coefficients.[4] In addition, we used the constrained solution which has the minimized values in AIC, BIC and CAIC.

Table 3 shows the goodness of fit statistics for model selection (ln-likelihood (LnL), AIC, BIC, CAIC) and the entropy measurement (EN) described above for K = 1, 2, 3, 4, 5 and 6. In the last section we mentioned the problems associated with locally optimum solutions. Therefore, we calculated all class options (2 through 6) ten times with different random starting values to be sure that we find a global maximum. For $K = 2$ the algorithm always found the same solution. For $K = 3$ we obtained different solutions close together, but the solution with the values shown in Table 3 was the one with the optimal goodness of fit measures (the smallest values for AIC, BIC and CAIC). The solutions for $K = 4$, $K = 5$ and $K = 6$ showed greater

---

[4] We note that negative coefficients are not "non-interpretable" per se. Negative values could indicate interaction effects involving the drivers of satisfaction. A disaggregate solution with negative values and a global optimum might cause the researcher to think about integrating interaction effects in the model. However, this is not the focus of our research. Therefore we use the model with the assumption of independent drivers of satisfaction.

variance. It shows that finding a global maximum becomes more difficult if we are looking for solutions with higher number of classes K. A reason for this phenomenon is the high-dimensional solution space for the local and global maxima. The success of iterative processes like the EM-algorithm depends on a good (plausible) set of starting values. As we do not know any good (or better plausible) sets of starting values, we have to increase the number of alternative solutions. Therefore for $K = 4$, $K = 5$ and $K = 6$ we started the approach twenty instead of ten times. Table 3 shows the solution with the minimized AIC, BIC and CAIC for $K = 4$, $K = 5$ and $K = 6$.

It is important to emphasize that the approach can be used in either an explorative or confirmative fashion. If the researcher knows any a priori information about the real values of the model, he or she can integrate this information when setting the starting values. If the algorithm finds a solution that corresponds to or is very similar to the starting values, it is evident that the prior information is a good start. In contrast, if the researcher wants to find new information about his or her model a number of different starting values are used to be sure that the result is not a local optimum in the high dimensional solution space. We focus on the 5-class solution where the AIC, BIC and CAIC measures are minimized. This solution also has the best entropy measure among the $K = 2$ through 6 solutions (EN = 0.43). Note that the selection of the most interpretable solution is the same as for the unconstrained case mentioned earlier: the 5-class solution.

Table 4 presents the path coefficients (impacts) for each of the 5 classes, where each class represents a relatively homogeneous group of customers. Going forward, we refer to these classes as market segments. Segments one through five comprise 10.7%, 36.8%, 17.2%, 27.7% and 7.6% of the overall survey population respectively. For segment one, satisfaction is almost

synonymous with safety, which has an impact of 0.984. The next highest and only other significant driver is cleanliness with an impact of 0.291. Segment two, the largest segment, is more balanced in that service, prices, cleanliness, convenience and safety all have significant impacts (ranging from 0.185 to 0.260). This segment also shows the largest impact of satisfaction on loyalty (0.863). Segment three is quite different from either of the previous two segments, as store layout and separate take out are the main drivers with impacts of 0.495 and 0.485. Segment four, which is the second largest overall, is the most price sensitive segment where impact of price is 0.210. Store layout and separate take out also have significant impacts for this segment. Segment five is marked by the importance placed on store layout and convenience, with impacts of 0.595 and 0.312 respectively. These shoppers want to find what they need and get in and out of the store quickly.

- insert Table 4 about here -

A membership probability is calculated for each customer in each segment. The entropy measure EN = .43 for $K = 5$ gives an aggregate value of how strongly customers belong to one particular segment. However the entropy measure gives no idea as to just what that means for each segment. For example, on the one hand, a customer can belong to four different segments with a membership probability of say .10 and to one segment with .60. On the other hand, a customer can belong to each segment with a membership probability of .20. Table 5 shows the number of customers who belong to a segment with a membership probability higher than .80, .60, .50 and .40 respectively. In our application only 130 customers belong to one segment with a membership probability higher than .80. This is about 13% of the whole sample. Ideally, membership probabilities should be as unique as possible for one specific segment, hence the probability should be near 1. But in reality, the lower membership probabilities illustrate the

complexity of measuring response-based variables. Table 5 shows that 686 customers out of 1,025 belong to one segment with a probability higher than .50. This means that our 5-segment solution is a fairly good approximation for grouping 1,025 different individuals together into 5 segments.

- insert Table 5 about here -

Post hoc analyses of the segments

To augment our interpretation of the segment-level results, we conducted post hoc analyses of the posterior probabilities of membership based on a model from Ramaswamy et al. (1993):

$$Q_{ik} = \sum_u Z_{iu} \delta_{uk} + V_{ik} , \qquad (21)$$

with

$$Q_{ik} = \ln(P_{ik} / P_i) ,$$

$$P_i = \left(\prod_k P_{ik} \right)^{1/K} \text{ as the geometric mean of the posterior probabilities,}$$

Z_{iu} as the value of descriptive variable u for individual i,

δ_{uk} as the impact coefficient for variable u for segment k,

V_{ik} as a random normal disturbance variable.

The descriptive variables in our study, collected as part of the convenience store survey, are: gender (male/female), age (in years), number of household members (5 categories), user frequency (daily, weekly, occasional user), education (3 categories), income (3 categories), 7-11 store user (Yes/No), neighborhood store user (Yes/No). The 7-11 brand was by far the most frequently measured convenience store in the sample, hence its use as a descriptive variable.

Also common were neighborhood store users who, when asked "At which convenience store do you shop most often?" they respond with a store name that is not part of a franchise system. This variable picks up the unique nature of "Mom and Pop" stores (typically family owned) that make up a large proportion of the industry. Table 6 shows the impact coefficients from the post hoc analysis of our 5-segment solution.

- insert Table 6 about here -

Overall there are relatively few significant descriptors for the five segments. One exception is gender, which is significantly related to segments two through five. For segment one, which is the safety conscious segment, household size is the largest descriptor. The larger the family, the more concern there is over safety. This is logical as larger families have more children who run errands or meet friends at convenience stores. Segment two, which had the most significant drivers of satisfaction (dominated by cleanliness), are primarily females who visited 7-11 stores. Segment three shoppers, where store layout and separate take out food are the dominant drivers, are primarily females who where not weekly shoppers. Segment four, the price conscious segment, is marked only by the fact that it is more female. In contrast, the store layout and convenience segment (the "get me in and out quickly" segment) is predominantly male.

Clearly, our analysis demonstrates that the results of an aggregate satisfaction model can be very misleading. Aggregate analysis hides the existence of meaningful subset of customers that are more homogeneous in their satisfaction drivers. While some customers are dominantly concerned with safety, other customers' satisfaction is the result of convenience or price. It is also clear from our post hoc analysis that the segments can not be clearly identified using simple descriptive variables. This is natural, as segments do not exist at the level of descriptive variables but rather at the level of benefits, consequences and needs (Best 2000). Yet marketing managers

- 23 -

often require such variables to derive market action implication. Gender is clearly one variable that helps to differentiate at least one segment. User frequency could also be such a variable. Managers in the convenience store industry pay particular attention to daily, weekly and occasional users and how their needs differ. In the next section of the paper, we use these a priori groupings before the customer satisfaction model is calculated.

Disaggregate PLS results: A priori segmentation

Table 7 shows the PLS results for the a priori segmentation based on the daily, weekly and occasional user segments (n = 265, 300 and 436 respectively). The solutions show that, in each segment, there are five to seven significant drivers of satisfaction, none of which have particularly large impacts. Some notable differences are the increased importance of separate take out for daily users (who likely obtain more of their meals from the stores), the importance of service to weekly customers, and the importance of products to occasional users.

- insert Table 7 about here -

But the pattern of results for each segment is similar to what we found for the aggregate sample. The finite mixture-based segments show much more pronounced differences in satisfaction drivers across segments. This suggests that, while the user frequency groups are homogeneous with respect to usage, they are still quite heterogeneous in their satisfaction drivers. The solution for each group is still an aggregate of different coefficients for the drivers of satisfaction.

Disaggregate PLS results: A priori and Finite Mixture PLS segmentation

To illustrate this heterogeneity and show how the Finite Mixture PLS approach provides insight to an existing, a priori segmentation scheme, we applied the new approach to the daily users segment. This "heavy user" group is of obvious import to convenience stores and a major

focus of their marketing activity. But not all daily users are necessarily looking for the same things from their convenience store. When we applied our new approach to the daily users, a two-segment solution emerged based on minimal values for the ln-likelihood, AIB, BIC and CAIC statistics (entropy = 0.60). The results are shown in Table 8.

- insert Table 8 about here -

Satisfaction for Segment 1 ($K = 1$, 22.8% of the sample) is driven dominantly by store layout and separate take out food, followed by motorist services. These "daily shoppers" fill their grocery baskets, stomachs and vehicles at their local convenience stores. They also appreciate high quality service. In contrast, segment 2 customers ($K = 2$, 77.2% of the sample) are more sensitive to safety, prices and cleanliness. These "daily stoppers" seem to stop to get just what they need. Satisfaction also has much more impact on loyalty for the segment 1 "shoppers" (0.823) than for the segment 2 "stoppers" (0.460). We applied our posthoc analysis approach described above to the two-segment solution one main difference emerged. Segment 1 customers are significantly more likely to shop at 7-11 stores.

Discussion and Conclusions

Satisfaction modeling has emerged as an important tool to help managers set priorities for improving quality, satisfaction and loyalty. However, satisfaction drivers should depend on the market segment of interest. Aggregate drivers reflect what is important for an "average" customer that may not exist. Clearly, satisfaction modeling should account for the heterogeneity in drivers that exists across segments. An emergent solution to capturing heterogeneity in market response is to use a latent class approach such as a finite mixture model. But whereas the latent class methods are based on maximum likelihood estimation, the operationalization of a satisfaction model often necessitates a least squares-based procedure. As a causal modeling

methodology, PLS (partial least squares) is particularly well suited to estimating and operationalizing satisfaction models in practice. PLS accommodates the skewed data and small sample sizes common in satisfaction research and, compared to other techniques, it is less prone to non-convergent or improper solutions. For managers, the performance scores and impacts that emerge from PLS analysis provide the diagnostic and benchmark information required to set priorities for improvement.

The goal of this article has been to merge the advantages of least square estimation, when estimating a satisfaction model, with the advantages of maximum likelihood estimation, when deriving market segments. Our Finite Mixture PLS approach is designed to capture heterogeneity in structural equation models that link quality and price drivers to satisfaction and subsequent loyalty. It empirically derives segments and directly estimates model relationships. The advantage of the approach compared to an a priori segmentation scheme is that the derived segments are homogenous in terms of model relationships. The approach calculates segment proportions, or the degree to which customers belong to particular segments, and the results can be statistically tested with goodness of fit measures. Thus the proposed Finite Mixture PLS model expands the existing Partial Least Squares approach to include one of the central issues in marketing theory and practice - segmentation.

When we apply the Finite Mixture PLS analysis to a national survey of quality, satisfaction and loyalty for convenience store customers, it reveals significant heterogeneity. Our five-segment solution identifies clear differences among customers who, for example, either value safety, separate take out and store layout, or prices. Another interesting observation is that, when we conduct a post hoc analysis that related descriptive variables to the segments, relatively few significant predictors emerge. Exceptions include gender, household size and frequency of

- 26 -

usage. This finding is consistent with the prevailing view in marketing that segments exist at the level of benefits, consequences and needs, while descriptive variables such as age, gender and frequency of use may be weak proxies (Best 2000). To illustrate the problem, we analyzed a prominent a priori segment, daily users, using the Finite Mixture PLS approach. The results again reveal clear differences in both the drivers of satisfaction and the effect of satisfaction on loyalty. Whereas "daily shoppers" value store layout, separate take out food and motorist services, "daily stoppers" value safety, prices and cleanliness.

Our findings reinforce an underlying premise in marketing that is often lost in practice, particularly in the practice of measuring and managing customer satisfaction. Satisfaction studies often rely on concrete, descriptive attributes of the product, service and customer segment. According to our findings and in line with means end theory, customers do not purchase a package of attributes, but rather a complex of benefits or even a set of values. And the benefit segments themselves are not easily described using traditional demographic variables. Applied satisfaction models should strive to capture both the abstract nature of satisfaction drivers and satisfaction-based market segments. Finite mixture-based segments that are built upon a latent variable modeling approach, such as PLS, can go a long way toward explaining variance in satisfaction judgements. They also help companies to draw more reasonable conclusions than those based on descriptive variables alone, such as frequency of usage.

One limitation of the proposed approach (mentioned in Footnote 4) is that it does not consider interaction effects in the inner model. In addition, in following the standard assumptions of the PLS approach, we assume that the regressions of the inner model are independent. Future research should focus on these aspects and on large-scale simulation studies to test the Finite Mixture PLS method in different marketing applications where heterogeneity is present.

Table 1

LATENT AND MEASUREMENT VARIABLES

Latent Variables	Measurement Variables
Service	Accuracy of the checkout Friendliness of the employees Attentiveness of the employees Grooming and appearance of the employees
Products	Stock or availability of products Brand names of products Variety and selection of products Freshness of products
Store layout	Ability to find what you need Neatness and orderliness of displays Feeling or sense of fun you get when walking through the store
Prices	Overall value Competitiveness of gasoline prices Competitiveness of store prices Frequency of sale items
Cleanliness	Cleanliness inside the store Cleanliness outside the store Cleanliness of the rest rooms
Convenience	Convenience of the location Hours of operation Speed and efficiency of employees Availability of parking
Safety	Lighting of the premises Ability to see what is happening in the store Feeling of safety and security
Motorist services	Accuracy of signs, gauges, and meters at the gas pumps Ability to pay at the pump Availability of air and water for vehicles Working operation of equipment such as gas pumps and air
Separate takeout (food)	Accuracy of food preparation Quality of takeout food
Customer Satisfaction	Overall satisfaction Overall performance versus expectations Overall performance versus an ideal convenience store
Customer Loyalty	Likelihood of revisiting Likelihood of recommending

Table 2
AGGREGATE RESULTS

		Endogenous Variables	
		Satisfaction	Loyalty
Endogenous Variables	Satisfaction		.625 (25.592)
	Loyalty		
Exogenous Variables	Service	.081 (2.940)	
	Products	.039 (1.483)	
	Store Layout	.174 (5.914)	
	Prices	.168 (7.071)	
	Cleanliness	.109 (3.728)	
	Convenience	.120 (4.322)	
	Safety	.193 (7.356)	
	Motorist Services	.039 (1.794)	
	Separate Take Out	.152 (6.560)	

Values in brackets are the t-statistics.

Table 3
MODEL SELECTION

Model	LnL	AIC	BIC	CAIC	EN
Aggregate (k = 1)	-2152.430	4305.860	4308.3263	4308.8263	1
k = 2	-2046.217	4096.434	4106.299	4108.299	.32
k = 3	-2018.076	4043.153	4060.416	4063.916	.41
k = 4	-1989.551	3989.103	4013.765	4018.765	.39
k = 5	-1951.910	3916.820	3948.881	3955.381	.43
k = 6	-1951.204	3918.409	3957.868	3965.868	.37

Table 4
DISAGGREGATE RESULTS: FIVE-SEGMENT SOLUTION

Segment		k = 1		k = 2		k = 3		k = 4		k = 5	
Segment Size		10.7%		36.8%		17.2%		27.7%		7.6%	
		Satisfaction	Loyalty	Satisfaction	Loyalty	Satisfaction	Loyalty	Satisfaction	Loyalty	Satisfaction	Loyalty
Endogenous Variables	Satisfaction		.551 (19.732)		.863 (16.881)		.532 (8.598)		530 (6.442)		.554 (20.077)
	Loyalty										
Exogenous Variables	Service	0*		.185 (4.719)		0*		.146 (1.766)		.046 (1.891)	
	Products	.104 (1.317)		019 (.538)		.180 (2.634)		0*		0*	
	Store Layout	0*		0*		.495 (6.688)		.197 (2.090)		.595 (18.451)	
	Prices	.103 (1.417)		.186 (5.986)		0*		.210 (2.907)		.189 (7.651)	
	Cleanliness	.291 (4.320)		.260 (6.426)		.006 (.078)		0*		.000 (.004)	
	Convenience	008 (.129)		.188 (4.877)		.110 (1.695)		.115 (1.375)		.312 (9.693)	
	Safety	.984 (15.438)		.189 (5.482)		.023 (.396)		.112 (1.422)		.051 (1.692)	
	Motorist Services	020 (.317)		0*		.071 (1.104)		087 (1.412)		.079 (3.593)	
	Separate Take Out	0*		.030 (1.031)		.485 (9.543)		.160 (1.985)		002 (.085)	

The values in brackets are the t-statistics
* Values are constrained to zero

Table 5
MEMBERSHIP PROBABILITIES: FIVE-SEGMENT SOLUTION

Membership Probability	Segment 1	Segment 2	Segment 3	Segment 4	Segment 5	TOTAL	%
> .80	12	18	6	93	1	130	.13
> .60	27	220	30	133	24	434	.42
> .50	41	370	50	166	59	686	.67
> .40	68	486	91	229	93	967	.94
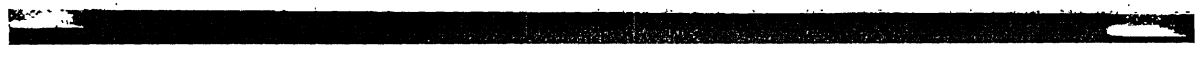
Table 6
POSTHOC ANALYSIS: FIVE-SEGMENT SOLUTION

		K = 1	K = 2	K = 3	K = 4	K = 5
Descriptive Variables	Gender	-.070 (-.197)	.769 (2.494)	.608 (2.036)	.855 (2.257)	-2.161 (-2.039)
	Age	.206 (.572)	-.374 (-1.201)	-.372 (-1.235)	-.524 (-1.372)	1.064 (.995)
	Household	.871 (2.419)	.004 (.014)	.073 (.241)	-.104 (-.271)	-.844 (-.788)
	Daily User	-.148 (-.139)	-.635 (-.691)	-.784 (-.881)	-.864 (-.764)	2.430 (.769)
	Weekly User	-.495 (-.454)	-1.385 (-1.465)	-1.726 (-1.884)	-1.966 (-1.692)	5.572 (1.714)
	Occasional User	-.763 (-.652)	-.694 (-.685)	-1.055 (-1.074)	-1.105 (-.888)	3.618 (1.038)
	Education high/less high	-1.833 (-.909)	-.631 (-.361)	-.949 (-.561)	-.478 (-.223)	3.891 (.648)
	Education tech/assoc	-2.634 (-1.254)	-.583 (-.320)	-1.005 (-.570)	-.525 (-.235)	4.746 (.759)
	Education coll/post_grad	-2.456 (-1.183)	-.541 (-.301)	-1.181 (-.678)	-.700 (-.317)	4.879 (.789)
	Income <=30	-.668 (-1.165)	-.118 (-.238)	-.048 (-.099)	-.262 (-.430)	1.096 (.642)
	Income >30 - <70	-.366) (-.631)	-.384 (-.766)	-.079 (-.162)	-.416 (-.675)	1.244 (.721)
	Income >=70	-.447 (-.825)	-.025 (-.053)	.221 (.485)	.005 (.009)	.246 (.153)
	7_11 Customer	-.004 (-.010)	.610 (1.946)	.490 (1.614)	.591 (1.534)	-1.686 (-1.565)
	Neigborhood Customer	-.295 (-.823)	.370 (1.190)	.257 (.853)	.340 (.890)	-.671 (-.628)
	Intercept	.340 (.983)	7.522 (25.141)	6.378 (21.992)	7.764 (21.113)	-22.004 (-21.380)

The values in brackets are the t-statistics

Table 7
A PRIORI SEGMENTATION BASED ON FREQUENCY OF USAGE

		Daily User n = 265		Weekly User n = 300		Occasional User n = 436	
		Satisfaction	Loyalty	Satisfaction	Loyalty	Satisfaction	Loyalty
Endogenous Variables	Satisfaction		.620 (12.869)		.599 (12.958)		.646 (17.660)
	Loyalty						
Exogenous Variables	Service	.058 (1.072)		.123 (2.491)		.073 (1.738)	
	Products	.064 (1.260)		.010 (.221)		.081 (1.976)	
	Store Layout	.153 (2.565)		.176 (3.463)		.186 (4.034)	
	Prices	.109 (2.227)		.182 (4.312)		.177 (4.846)	
	Cleanliness	.141 (2.327)		.121 (2.246)		.060 (1.354)	
	Convenience	.100 (1.768)		.162 (3.243)		.142 (3.372)	
	Safety	.150 (2.744)		.189 (3.886)		.214 (5.493)	
	Motorist Services	.070 (1.479)		.045 (1.060)		.046 (1.389)	
	Separate Take Out	.213 (4.351)		.083 (1.907)		.115 (3.306)	

Table 8
DAILY USERS: TWO-SEGMENT SOLUTION

		K = 1		K = 2	
		22.8%		77.2%	
		Satisfaction	Loyalty	Satisfaction	Loyalty
Endogenous Variables	Satisfaction		.823 (5.910)		.460 (12.985)
	Loyalty				
Exogenous Variables	Service	.211 (2.353)		0*	
	Products	.210 (1.670)		.015 (.268)	
	Store Layout	.448 (3.482)		.030 (.465)	
	Prices	.114 (1.081)		.170 (3.128)	
	Cleanliness	0*		.170 (2.362)	
	Convenience	0*		.120 (1.831)	
	Safety	.016 (.167)		.248 (3.638)	
	Motorist Services	.267 (2.485)		.063 (1.222)	
	Separate Take Out	.406 (3.604)		.167 (3.116)	

The values in brackets are the t-statistics
* Values are constrained to zero

Figure 1.

METHODOLOGICAL TAXOMONY FOR CAPTURING HETEROGENEITY IN
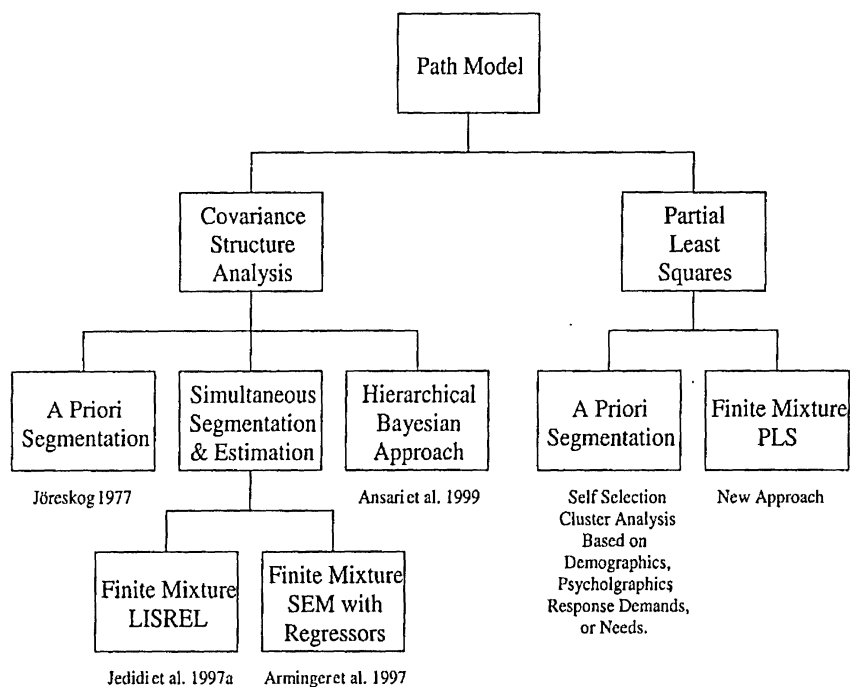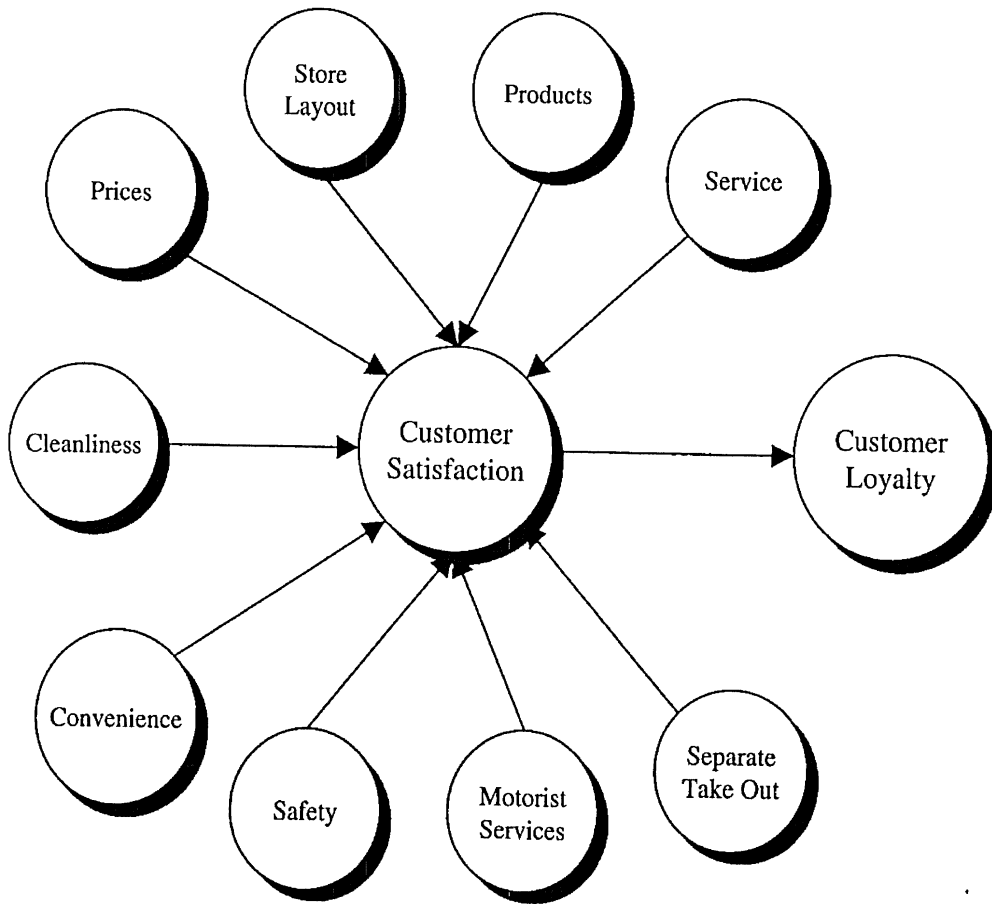
STRUCTURAL EQUATION MODELS

- 36 -

Figure 2.

THE NACS CUSTOMER SATISFACTION MODEL

References

Akaike, Hirotugu (1974), "A New Look at Statistical Model Identification," *IEEE Transactions on Automatic Control*, 6, 716-723.

Ansari, Asim, Kamel Jedidi and Harasharanjeet S. Jagpal (1999) "A General Hierarchical Bayesian Methodology for Treating Hetergogeneity in Structural Equation Models," unpublished paper.

Arminger, Gerhars and Petra Stein (1997), "Finite Mixtures of Covariance Structure Models With Regressors," *Sociological Methods & Research*, 26 (2), 148-182.

Bagozzi, Richard P. (1982), " A Field Investigation of Causal Relations Among Cognitions, Affects, Intentions, and Behavior, " *Journal of Marketing Research*, 19 (November), 562-584.

_____ (1994), "Structural Equation Models in Marketing Research: Basic Principles," in *Principles of Marketing Research*, Richard P. Bagozzi, ed. Cambridge, MA: Blackwell, 317-385.

_____ and Youjae Yi (1994) "Advanced Topics in Structural Equation Models," in *Advanced Methods of Marketing Research*, Richard P. Bagozzi, ed. Cambridge, MA: Blackwell, 1-52.

Best, Roger J. (2000), *Market-Based Management: Strategies for Growing Customer Value and Profitability*, Upper Saddle River, NJ: Prentice Hall.

Bozdogan, Hamparsum and Stanley L. Sclove (1984), "Multi-Sample Cluster Analysis Using Akaike's Information Criterion," *Annals of the Institute of Statistical Mathematics*, 36, 163-180.

Churchill, Gilbert A. and Carol Suprenant (1982), "An Investigation into the Determinants of

    Customer Satisfaction," *Journal of Marketing Research*, 19 (November), 491-504.

Day, Ralph L. (1977), "Extending the Concept of Consumer Satisfaction," in *Advances in*

    Consumer Research, Vol. 4, William D. Perreault, ed. Atlanta: Association for Consumer

    Research, 149-154.

Dempster, A.P, N.M. Laird and D.B. Rubin (1977), "Maximum likelihood from incomplete data

    via the EM-algorithm," *Journal of the Royal Statistical Society: Series B*, 39, 1-38.

DeSarbo, Wayne S. and William L. Cron (1988), "A Maximum Likelihood Methodology for

    Clusterwise Linear Regression," *Journal of Classification*, 5, 249-282.

Fornell, Claes (1987), "A Second Generation of Multivariate Analysis: Classification of Methods

    and Implications for Marketing Research," in *Review of Marketing 1987*, Michael J.

    Houston, ed. Chicago: American Marketing Association.

_____ (1995), "The Quality of Economic Output: Empirical Generalizations About Its

    Distribution and Association to Market Share," *Marketing Science*, 14 (Summer), G203-

    G211.

_____ and Fred L. Bookstein (1982), "Two Structural Equation Models: LISREL and PLS

    Applied to Consumer Exit-Voice Theory," *Journal of Marketing Research*, 14

    (November), 440-452.

_____ and Jaesung Cha (1994), "Partial Least Squares," in *Advanced Methods of*

    Marketing Research, Richard P. Bagozzi, ed. Cambridge, MA: Blackwell, 52-78.

_____, Michael D. Johnson, Eugene W. Anderson, Jaesung Cha and Barbara Everitt

    Bryant (1996), "The American Customer Satisfaction Index: Nature, Purpose and

    Findings," *Journal of Marketing*, 60 (October), 7-18.

Gustafsson, Anders and Michael D. Johnson (1997), "Bridging the Quality-Satisfaction Gap,"

    Quality Management Journal, 4 (3), 27-43.

Heskett, James L., Sasser, W. Earl and Schlesinger, Leonard A. (1997), *The Service Profit*

    Chain, New York, The Free Press.

Jedidi, Kamel, Harasharanjeet S. Jagpal and Wayne S. DeSarbo (1997a), "Finite-Mixture

    Structural Equation Models for Response-Based Segmentation and Unobserved

    Heterogeneity," *Marketing Science*, 16 (Winter), 39-59.

Jedidi, Kamel, Harasharanjeet S. Jagpal and Wayne S. DeSarbo (1997b), "STEMM : A General

    Finite Mixture Structural Equation Model," *Journal of Classification*, 14 (1), 23-50.

Jedidi, Kamel, Venkatram Ramaswamy, Wayne S. DeSarbo and Michel Wedel (1996), "On

    Estimating Finite Mixtures Of Multivariate Regression and Simultaneous Equation

    Models," *Structural Equation Modeling*, 3 (3), 266-289.

Johnson, Michael D. and Anders Gustafsson (2000), *Improving Customer Satisfaction, Loyalty*

    and Profit: An Integrated Measurement and Management System, San Francisco, CA:

    Jossey-Bass.

Jöreskog, Karl G. (1977), "Structural Equation Models in the Social Sciences: Specification,

    Estimation, and Testing," in *Applications of Statistics*, P. K. Krishnaiah, ed. Amsterdam:

    North-Holland, 265-287.

Martilla, John A., and John C. James (1977), "Importance-Performance Analysis," *Journal of*

    Marketing, 41 (January), 77-79.

McLachlan, Geoffrey J. and Kaye E. Basford (1988), *Mixture Models: Inference and*

    Aapplications to Clustering. New York: Marcel Dekker.

McLachlan, Geoffrey J. and Thriyambakam Krishnan (1997), *The EM-Algorithm and Extensions*. New York, Wiley.

Muthén, Bengt O. (1989), "Latent Variable Modeling in Heterogenous Populations," *Psychometrika*, 54, 557-585.

Oliver, Richard L. (1997), Satisfaction: A Behavioral Perspective on the Consumer. New York: McGraw-Hill.

Ramaswamy, Venkatram, Wayne S. DeSarbo, David J. Reibstein and William T. Robinson (1993), "An Empirical Pooling For Estimating Marketing Mix Elasticities With PIMS Data," *Marketing Science*, 12 (Winter), 103-124.

Schwarz, G. (1978), "Estimating the Dimension of a Model," *Annals of Statistics*, 6, 461-464.

Steenkamp, Jan-Benedict E. M. and Hans C. M. van Trijp (1996), "Quality Guidance: A Consumer-Based Approach to Food Quality Improvement using Partial Least Squares", *European Review of Agricultural Economics*, 23, 195-215.

Wedel, Michel and Wagner A. Kamakura (1999), *Market Segmentation, Conceptual and Methodological Foundations*, Second Edition, Boston, Dortrecht, London: Kluwer Academic Publishers.

Wold, Hermann (1966), "Estimation of Principal Components and Related Models by Iterative Least Squares," in *Multivariate Analysis: Proceedings of an International Symposium Held in Dayton, Ohio*, P.R. Krishnaiah, ed. New York: Academic Press. 391-420.