Marketing scholars commonly characterize market structure by studying the patterns of substitution implied by brand switching. Though the approach is useful, it typically ignores the destabilizing role of marketing variables (e.g., price) in switching behavior. The authors propose a flexible choice model that partitions the market into consumer segments differing in both brand preference and price sensitivity. The result is a unified description of market structure that links the pattern of brand switching to the magnitudes of own- and cross-price elasticities. The approach is applied in a study of competition between national brands and private labels in one product category.

**A Probabilistic Choice Model for Market Segmentation and Elasticity Structure**

One of the most useful concepts in marketing is consumer segmentation. Numerous bases for segmentation can be advanced, each with its own set of advantages and disadvantages for particular types of product policy issues (Wind 1978). However, as Blattberg et al. (1978) point out, the managerial relevance of a segmentation procedure is related directly to its ability to partition the consumer population into relatively homogeneous groups that differ substantially in purchase behavior.

Recently, Grover and Srinivasan (1987) proposed that brand choice probabilities be used to define both market segments and market structure. Briefly stated, market structure is the classification of brands into submarkets that have a higher degree of competition than the market taken as a whole (Day, Shocker, and Srivastava 1979). The key idea behind the Grover-Srinivasan approach is that the same choice probabilities that provide a basis for behavioral segmentation also generate patterns of brand switching that reveal the structure of a product market. These researchers demonstrate how latent class analysis (see, e.g., Bartholomew 1987) applied to a brand switching matrix can be used to infer the types of preference segments composing a given market.

An important assumption required by this approach (and many others) is the presumed stability of the market. In particular, the analyst must assume that the major determinant of observed brand switching is the fundamental stochasticity in choice (Bass 1974), not temporal variation in elements of the marketing mix. This restrictive assumption can be relaxed by adding more complexity to the model. For example, by applying the Lancasterian choice framework of the Defender model (Hauser and Shugan 1983), Shugan (1987) demonstrates that temporal changes in brand prices and market shares can be used to calibrate simultaneously the location of brands in a perceptual space (i.e., the market structure) and the relative sizes of various preference segments. This approach is attractive because it combines market structure, consumer segmentation, and price sensitivity in a common framework.

The purpose of our research is to propose and evaluate a new approach to preference segmentation that enables the researcher to identify the underlying determinants of brand switching probabilities and aggregate response to price changes. In contrast to previous approaches, the segmentation model is calibrated with household-level data and then the patterns of brand substitutability likely to be observed in aggregate data are inferred. Because we allow for nonstationarity due to price variation, we obtain a representation of market structure that simultaneously reveals the brand preferences of key consumer
segments and enables us to predict the magnitude of aggregate own- and cross-price elasticities. Thus, the model provides a conceptual link between the analysis of brand switching (Grover and Srinivasan 1987; Kalwani and Morrison 1977; Urban, Johnson, and Hauser 1984) and the decomposition of aggregate brand price elasticities (Cooper 1988; Russell and Bolton 1988).

We first specify the model and develop a simple estimation procedure. We also show how these measures of segment preferences and price sensitivities can be used to develop an alternate representation of market structure in terms of price elasticities. We then apply the theory in an empirical study of the competitive structure linking national brands and private labels. Our results are consistent with the notion that consumer perceptions of a price-quality relationship create a distinctive structure in both preferences and elasticities. We conclude with suggestions for further research.

A MODEL FOR PREFERENCE SEGMENTATION

Our approach rests on the assumption that consumers can be placed into a small number of segments, each characterized by a vector of mean preferences and a single price sensitivity parameter. For such a model to be useful, we must be able both to calibrate the segment-level parameters and to determine the likely segment membership of particular consumers. We consider these issues in turn.

Random Utility Theory

In deriving the model, we start with the usual assumptions of random utility theory (e.g., Currim 1982; Kamakura and Srinavastava 1984): when facing a purchase decision, consumers assign random utilities to each brand considered and then select the one with the highest derived utility. We decompose this utility into a deterministic component, which depends on the intrinsic characteristics of the brand and its price (and/or other marketing mix variables), and a random component. Hence, the random utility assigned to brand \( j \) by consumer \( k \) at purchase occasion \( t \) is

\[
U_{jk} = u_{jk} + \beta_k X_{jk} + \epsilon_{jk},
\]

where \( u_{jk} \) is the intrinsic utility/value of brand \( j \) for consumer \( k \), \( \beta_k \) is the price parameter for consumer \( k \), \( X_{jk} \) is the net available price of brand \( j \) for consumer \( k \) at time \( t \), and \( \epsilon_{jk} \) is a random error.

We further assume that the stochastic components \( \epsilon_{jk} \) are independent, identically distributed Weibull. Thus, the conditional probability of choosing brand \( j \) at time \( t \) is given by the multinomial logit model,

\[
P_j(u, \beta, X_{jt}) = \exp(u_{j\beta} + \beta_j X_{jt}) / \sum_j \exp(u_{j\beta} + \beta_j X_{jt})
\]

Preference Segmentation

Rather than estimate subject-specific parameters \( u_{jk} \) and \( \beta_k \), we assume the existence of \( i = 1, 2, \ldots, M \) homogeneous segments with relative sizes \( f_i = \exp(\lambda_i) / \sum_i \exp(\lambda_i) \).

In principle, by making \( M \) sufficiently large, it is possible to explain all variability in preferences and price sensitivity. In practice, we prefer parsimony and attempt to represent the market using a small value for \( M \). We show subsequently how to relax the homogeneity assumption to achieve this goal.

The key to our model is expressing a consumer’s choice probabilities in terms of the choice probabilities corresponding to the various segments. By relabeling the parameters of equation 2, we write the probability of choosing brand \( j \), conditional on consumer \( k \) being a member of segment \( i \), as

\[
P_j(u, \beta, X_{jt}) = \exp(u_{ji} + \beta_{ji} X_{jt}) / \sum_j \exp(u_{ji} + \beta_{ji} X_{jt})
\]

Because \( f_i \) represents the likelihood of finding a consumer in segment \( i \), the unconditional probability of choice for brand \( j \) by consumer \( k \) can be computed as

\[
P_j(u, \beta, X_{jt}) = \sum_i f_i P_j(u, \beta, X_{jt})
\]

Following the extensive literature on latent variable models (see, e.g., Bartholomew 1987; Dill and Mullani 1989), we assume that the unconditional choice probabilities can be decomposed into a weighted average of underlying (or “latent”) choice probabilities. We interpret this decomposition as a representation of the market’s preference segmentation. Thus, \( f_i \), the likelihood of finding a consumer in segment \( i \), is viewed as the relative size of the segment in the population of consumers.

Estimation

As demonstrated in equation 5, the segment-level parameters can be estimated by using the unconditional probability \( P_j(u, \beta, X_{jt}) \) to infer both the relative segment sizes \( f_i \) and the segment probabilities \( P_j(u, \beta, X_{jt}) \). Let us consider the choice history of consumer \( k \) during a time interval \( T \),

\[
H_k = c(t), \quad t = 1, 2, \ldots, T,
\]

where \( c(t) \) is the index of the chosen brand at occasion

\[1\] Heterogeneity in consumer choice rules can also be modeled by assuming that the parameters of equation 1 follow a multivariate normal distribution across the population of consumers (Elrod 1988, Hausman and Wise 1978, Kamakura and Srinavastava 1986) This assumption implies a unimodal preference distribution and consequently ignores the possibility of preference segments. In theory, the logit formulation for each segment could be replaced by a random coefficients model. However, the added complexity of the probit formulation would render the model infeasible in practice.
t. The likelihood of this choice history can be computed as

\[ L(H_t) = \sum_i \left[ \exp(\lambda_i) L(H_t|t) / \sum_j \exp(\lambda_j) \right] \]

where:

\[ L(H_t|t) = \prod_i \mathbb{P}_{ct}(u_i, \beta_i, X_{t_i}). \]

The preceding expression for the likelihood of an observed choice history assumes independence among choice decisions made by one consumer over all choice occasions. This assumption of a zero-order choice process is commonly used in market structure models (e.g., Grover and Srinivasan 1987) and is the foundation of the extensive literature on stochastic brand choice (e.g., Bass 1974). However, models based on aggregate data also need the additional assumption of stationarity (i.e., choice probabilities are constant over the sampled time period). By including price as an exogenous variable at the consumer level, we allow for nonstationarity due to price fluctuations and are able to calibrate its influence.

Maximum likelihood estimates of the model parameters can be obtained in a straightforward way by using the choice histories from a sample of consumers and the likelihood expression in equation 6. Details are provided in Appendix A. The algorithm leads to estimates of the price sensitivity \( \beta \), and mean brand utilities \( u_i \) for each segment \( i \) in the market and the probabilities of randomly drawing a consumer from each segment \( f_i \).

Because the number of segments \( M \) is unknown, parameter estimation is carried out conditional on an assumed value for \( M \). In practice, the number of segments is selected to minimize a variant of Akaike’s information criterion (Judge et al. 1980, p. 423). We elaborate on this approach in the empirical work reported subsequently.

**Assigning Consumers to Segments**

The probability of membership in a particular segment \( i \), conditional on the observed choice history, is obtained by revising the prior probability of membership \( f_i \) in a Bayesian fashion.

\[ P(k \in \phi|H_t) = L(H_t|t) f_i / \sum_{i'} L(H_t|t') f_{i'}. \]

We use these posterior probabilities in our empirical study.

**REPRESENTING COMPETITIVE STRUCTURE**

The proposed model enables the analyst to describe the competitive brand structure in two ways. First, by inserting the estimated segment-level parameters and the average brand prices \( X_{t}, \) into equations 3 and 4, we can predict both the within-segment brand choice shares and the relative size of the segments. This representation of preference segmentation is similar to the market structure obtained by using the Grover-Srinivasan (1987) approach on brand switching data.

Second, we can analyze the competitive structure in terms of price sensitivity. A distinctive feature of our model is that it estimates the “intrinsic preference” \( u_i \) for each brand by a market segment while sorting out consumers’ reactions to price differentials. This feature enables us to determine whether choice shares within a segment are due to preferences for each brand or to price sensitivity. It also provides a straightforward means of constructing a market-level matrix of cross-price elasticities. We next elaborate on this alternate representation of market structure.

**Deriving Aggregate and Segment-Level Share Elasticities**

Given the logit formulation at the segment level and our assumption of homogeneity within segments, the choice share (own and cross) elasticities \( \eta_{i,j} = \frac{X_{j} / S_{i}}{\partial S_{i} / \partial X_{j}} \) within a segment \( i \) are given by (e.g., Russell and Bolton 1988)

\[ \eta_{i,j} = \beta (1 - S_{j}) X_{j}, \]

\[ \eta_{i,j} = -\beta S_{j} X_{j}, \quad j \neq j', \]

where:

\[ S_{i} = \exp(u_{i} + \beta X_{i}) / \sum_{i'} \exp(u_{i'} + \beta X_{i'}) \]

is the aggregate share of brand \( j \) within segment \( i \).

The aggregate (over all segments) cross elasticities are obtained by combining the segment-level elasticities,

\[ \eta'_{i,j} = \left( X_{i} / \sum_{i'} f_{i} S_{i} \right) \left( \sum_{i} f_{i} \partial S_{i} / \partial X_{j} \right) \]

\[ = \left( X_{i} / \sum_{i} f_{i} S_{i} \right) \left( \sum_{i} f_{i} (-\beta S_{j} X_{j}) \right) \]

\[ = \left[ \sum_{i} (f_{i} S_{i})(-\beta S_{j} X_{j}) \right] / \left[ \sum_{i} f_{i} S_{i} \right] \]

\[ = \sum_{i} f_{i} S_{i} / S_{j} \eta'_{i,j}. \]
where $S_j = \sum_i f_i S_j$ is the overall choice share of brand $j$. That is, the aggregate elasticities are weighted averages of the segment-level elasticities. Using similar arguments, we can show that the same relation holds for the aggregate own-price elasticities,

$$\eta_0 = \sum_i [(f_i S_j)/S_j] \eta_{0i}$$

These aggregate elasticities portray the competitive structure of the marketplace, showing the impact of price changes in a brand on its competitors and its vulnerability to competitive pricing strategies. Aside from the overall assessment of competitiveness among the brands, our model provides detailed information on the competitive structure at the segment level (i.e., segment-level elasticities and shares and segment sizes), thus offering some clues on the reasons for the market impact and vulnerability of a brand.

**Correction for Heterogeneity Bias**

Though our model is based on the assumption of homogeneous segments, we expect some degree of heterogeneity to remain in any parsimonious representation of the preference segmentation. Because in the estimates in equations 9 and 10 we assume homogeneity, the predicted elasticities will be biased when the utility parameters differ substantially across consumers (Guadagni and Little 1983). A correction for this bias is discussed next.

Define $S_j^{(i)}$ as the share of choices for brand $j$ by consumer $k$. Assume that within each segment the value of a consumer’s $\beta_j$ is a poor predictor of his or her preferences $S_j^{(i)}$ for any set of given prices. Then, regardless of the degree of preference heterogeneity within a segment, the true segment-level elasticities can be written as

$$\eta_{0i} = \gamma_{0i} \eta_{0i},$$

$$\eta_{ij} = \gamma_{ij} \eta_{ij}, \quad j \neq j',$$

where $\eta_{0i}$ and $\eta_{ij}$ are defined as in equations 9 and 10,

$$\gamma_{0i} = [S_{0i} - E(S_{0i} S_{ij})]/[S_{0i} - S_{ij}],$$

$$\gamma_{ij} = E(S_{0i} S_{ij})/S_{0i},$$

and $E(\cdot)$ denotes a mean taken over all consumers in segment $i$. Details are provided in Appendix B.

Clearly, equations 15 and 16 imply that $\eta_{0i}^* = \gamma_{0i} \eta_{0i}$ and $\eta_{ij}^* = \gamma_{ij} \eta_{ij}$ whenever the segment is homogeneous. However, in general, estimates of a segment’s own-price elasticities $(X_j, S_j) \partial S_j/\partial X_j$ are too large when heterogeneity is ignored.

In the empirical work reported next, we first calculate unadjusted segment-level elasticities $(\eta_{0i}^* \text{ and } \eta_{ij}^*)$ using equations 9 and 10 and then compute adjusted elasticities $(\eta_{0i}^* \text{ and } \eta_{ij}^*)$ using equations 13 through 16. Market-level elasticities are estimated by inserting $\gamma_{0i}$ and $\gamma_{ij}$ (in place of $\eta_{0i}$ and $\eta_{ij}$) in equations 11 and 12. As consumers are not deterministically assigned to one segment, sample estimates of $E(S_j^{(i)} S_{ij}^{(i)})$ for segment $i$ must be weighted by the posterior probability of segment membership defined in equation 8. That is,

$$E(S_j^{(i)} S_{ij}^{(i)}) = \frac{\sum_j S_j^{(i)} S_{ij}^{(i)} P(k \in i | H_i)}{\sum_k P(k \in i | H_i)}$$

where the summation is over all consumers in the market.

If a more restrictive assumption is made—that the choice shares $S_j^{(i)}$ are Dirichlet-distributed within segment $i$ (Jeuiland, Bass, and Wright 1980) with heterogeneity parameter $\rho_i (0 \leq \rho_i \leq 1)$—then

$$E(S_j^{(i)} S_{ij}^{(i)}) = \rho_i S_{0i} (1 - \rho_i) S_{ij}$$

$$E(S_j^{(i)} S_{ij}^{(i)}) = (1 - \rho_i) S_{0i} S_{ij}$$

where the expectations are over all consumers in segment $i$. Inserting these definitions in equations 15 and 16, we obtain the interesting result that $\eta_{0i} = \gamma_{0i} = (1 - \rho_i)$. That is, within a segment, all correction factors are equal. Though our empirical work is not based on the Dirichlet assumption, we show subsequently that the correction factors are relatively constant within each segment.

**APPLICATION OF THE MODEL**

To illustrate the approach, we analyze the competition among national brands and private labels in one product category. Our purpose is to show how the model’s estimates of brand preferences and price sensitivities can be used to construct a managerially useful description of brand competition. However, this example is chosen also to enable us to explore the characteristics of competition between national brands and private labels. Our analysis shows that this market conforms to a theory of asymmetric price-tier competition proposed by Blattberg and Wisniewski (1985).

**Description of Data**

The data analyzed here consist of 78 weeks of retail scanner data collected by Information Resources, Inc.

---

4 More precisely, we assume that within a segment and conditional on any set of prices, the distributions of $\beta_j$ and $S_j^{(i)}$ are independent. Though this assumption is not generally valid for an entire market (e.g., price-sensitive consumers may have higher purchase probabilities for private label brands), it may be a good approximation within a relatively homogeneous segment. The condition always holds if we assume that the $\beta_j$'s of all consumers within a segment are identical. Further discussion of issues encountered in the aggregation of logit models is provided by Allenby and Rossi (1988).

5 Because $S_j^{(i)} \leq E(S_j^{(i)} S_{ij}^{(i)}) \leq S_{0i}$, it follows that $0 \leq \gamma_{0i} \leq 1$. Also, because $E(S_j^{(i)} S_{ij}^{(i)})$, $S_{0i}$, and $S_{ij}$ are non-negative, $\gamma_{0i} \leq 0$. However, no upper bound for $\gamma_{0i}$ is generally valid.
Table 1

COMPETITIVE BRAND STATISTICAL SUMMARY

<table>
<thead>
<tr>
<th>Brand name</th>
<th>Mean price</th>
<th>Loyal households</th>
<th>Switching households</th>
<th>All households</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Choice share</td>
<td>Volume share</td>
<td>Choice share</td>
</tr>
<tr>
<td>A</td>
<td>4.29</td>
<td>74.7</td>
<td>74.6</td>
<td>26.3</td>
</tr>
<tr>
<td>B</td>
<td>3.54</td>
<td>15.5</td>
<td>16.5</td>
<td>30.8</td>
</tr>
<tr>
<td>C</td>
<td>3.38</td>
<td>6.8</td>
<td>6.3</td>
<td>28.0</td>
</tr>
<tr>
<td>P</td>
<td>3.09</td>
<td>2.9</td>
<td>2.5</td>
<td>15.0</td>
</tr>
</tbody>
</table>

* Loyal households are defined as families buying only one brand during the 78-week period. They represent 31.5% of the total number of households and 19.9% of the total market volume.

Our data contain price and coupon information for the purchased brand. Competitive prices are inferred by examining the purchase histories of all households. If possible, we use information from the same store and week as for the purchased brand. If a brand is not sold in a given store and week, we infer its price by using the average shelf price for the brand across all stores within the week in which the purchase was made. In a few cases, this second definition also fails. In such instances, we use the average brand price taken over all stores and all weeks. This process of averaging to reconstruct the competitive price set is common in studies using scanner data (see, e.g., Tellis 1988).

**Definition of Loyalty**

Because our model cannot compute the relative sizes of loyalty segments directly, we separated completely loyal households from the rest prior to the analysis. We define loyalty as a purchase pattern in which various package sizes of only one brand are purchased during the 78-week observation period. These loyal households represent 31.5% of our sample. The remaining 68.5% are classified as switching households and are used to calibrate our model.

As shown in Table 1, brand preferences are markedly different across these two groups. Brand A, a premium-priced national brand that devotes considerable resources to national television advertising, constitutes 75% of the loyalty group. The remaining national brands (B and C) are better positioned for the switching group. Brand B supports its higher price with a relatively small amount of national advertising. Brand C competes primarily on price. Though other products carrying C's family brand name are advertised, C's producer does not specifically advertise in this category.

**Choosing the Number of Segments**

As noted before, we defined loyalty segments on the basis of a prior classification of households. The pur-
purchase histories of the remaining households (401 of the total 585) were used to calibrate the parameters of the preference segmentation model.

Model parameters were estimated by applying the maximum likelihood procedure described in Appendix A, conditional on an assumed number of switching segments. To specify the model in a reasonable way, we systematically varied M, the number of segments, and then calculated Akaike’s information criterion (Judge et al. 1980, p. 423),

\[ \text{AIC} = -2(LL - p)/N, \]

where \( LL \) is the maximum value of the log likelihood, \( p \) is the number of parameters (equal to \( 5M - 1 \) in this case), and \( N = 3615 \) is the total number of datapoints (i.e., number of purchase occasions). Minimizing the AIC leads to a model that does not overfit the data.

On the basis of the following results,

<table>
<thead>
<tr>
<th>( M )</th>
<th>( AIC )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2 351</td>
</tr>
<tr>
<td>2</td>
<td>2 184</td>
</tr>
<tr>
<td>3</td>
<td>2 151</td>
</tr>
<tr>
<td>4</td>
<td>2 066</td>
</tr>
<tr>
<td>5</td>
<td>2 038</td>
</tr>
<tr>
<td>6</td>
<td>2 036</td>
</tr>
</tbody>
</table>

we chose a five-segment representation of the switching households. Though the AIC is not quite minimized at this point, it does not change appreciably after five segments are extracted.

**Preference Segmentation**

The maximum likelihood estimates of the parameters for the five-segment solution are reported in Table 2. These parameters correspond to the model of equations 3 through 5: segment-level mean utilities for each brand \( u_i \), segment-level mean price sensitivity \( \beta \), and the segment size measured in relation to all switching households \( \lambda \). Because the parameters of the multinomial logit are uniquely determined up to an additive constant (Guadagni and Little 1983), certain parameters have been set to zero in the model calibration. These constraints have no impact on our interpretation of the preference segmentation.

One key conclusion is apparent by inspecting Table 2. With the exception of segment 5, all segments have a high intrinsic utility for A. Recalling the formula for the relative segment size \( f_i = \exp(\lambda_i)/\sum_j \exp(\lambda_j) \) and allowing for measurement error, we can infer that between 75% and 90% of the switching segment would select A as a first choice if all brand prices were identical. This finding reinforces the impression that A has created a high quality image through its pricing and advertising strategies.

A more useful summary of the preference segmentation can be obtained by using equation 4 and mean brand prices (Table 1) to transform the raw parameter estimates into average purchase probabilities. These probabilities, reported in Table 3, are supplemented with information based on the prior classification of loyal households. Because loyal households make up 31.5% of our entire sample, we normalized the switching segment proportions as \( f_i^* = (1 - .315)f_i \), thus forcing the sum of the switching segment proportions to represent 68.5% of the market. We also decomposed the loyal households into four groups by counting the number of households loyal to each brand.

The summary in Table 3 is similar to the market structure obtained by a Grover-Srinivasan (1987) analysis of brand-switching data. The important difference—a difference that we exploit subsequently—is calibration of the price sensitivities of the various switching segments. For example, we obtain the plausible result that the most price-sensitive group (segment 4) has a high purchase probability for private labels.

**Price Tier Structure**

The most striking aspect of the preference segmentation is the organization of segments 1 through 4 in terms of price tiers. For purposes of exposition, we have underlined the purchase probabilities greater than .1 in these four segments. Recalling that the brands are ordered in terms of price level (A high, P low), we see that in relation to a segment’s most preferred brand, consumers will switch “up” to more expensive brands but will not switch “down” to less expensive brands. For example, segment 3 is relatively price sensitive and prefers the national brand C. When this segment switches to less preferred brands, it almost always buys more expensive national brands; it rarely switches to private labels.

This type of asymmetric switching has been proposed by Blattberg and Wisniewski (1985) as a model of competition between national brands and private labels. Their argument is that asymmetric patterns of brand switching are caused by a perceived relationship between price and
quality. In effect, consumers will not buy below some self-defined quality level. Our empirical results indicate that the price tier effect can be found also within national brands—not just between national brands and private labels.

The clear anomaly in this structure is segment 5. Its price coefficient is statistically insignificant and its average preferences show no readily interpretable pattern. In comparing the four-segment solution (not shown) with Table 3, we discovered that the four-segment solution is virtually identical except that segments 3 and 5 are combined. Apparently, segment 5 represents the residual of all the switching segments once the price tier structure is removed.

**ELASTICITY STRUCTURE**

Using the relationships developed previously, we can also develop a representation of brand competition in terms of price elasticities. This representation complements the preference segmentation by revealing how segment differences in size, average preference, and inherent price sensitivity affect the market response to price changes.

**Correction Factors**

As argued before, despite the assumptions of our model, we do not expect each of the preference segments to be entirely homogeneous. In Table 4 we give the correction factors that must be used to obtain unbiased estimates of the own- and cross-price choice elasticities. Significantly, all correction factors are between zero and one and the within-segment variation in these factors is small. Both findings are consistent with the idea that the preferences within each segment are approximated by a Dirichlet distribution. Because a correction factor equal to one implies no bias in the unadjusted elasticities, our results imply that the heterogeneity bias is relatively small. Notice that the observed averages of the correction factors range from .76 to .82.

**Choice Share Elasticities**

In Table 5 we summarize the predicted choice share elasticities for segments 1 through 4 and for the market in total. Because switching segment 5 has a statistically insignificant β, we assume that its true elasticities are zero and do not provide a segment-level elasticity matrix. For each segment, the average segment-level choice probabilities and mean brand prices also are listed. For the total market, the predicted shares were obtained by weighting the choice probabilities of all segments in Table 3 by the relative segment sizes. It is interesting to

---

### Table 3

**PREFERENCE SEGMENTATION AND PRICE SENSITIVITY**

<table>
<thead>
<tr>
<th>Loyal segments</th>
<th>Switching segments*</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>P</td>
<td>1</td>
</tr>
<tr>
<td><strong>Choice probabilities</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>1</td>
<td>.790</td>
<td>219</td>
<td>.152</td>
<td>095</td>
</tr>
<tr>
<td>B</td>
<td>1</td>
<td>.089</td>
<td>646</td>
<td>.259</td>
<td>238</td>
</tr>
<tr>
<td>C</td>
<td>1</td>
<td>.069</td>
<td>092</td>
<td>.520</td>
<td>301</td>
</tr>
<tr>
<td>P</td>
<td>1</td>
<td>.052</td>
<td>043</td>
<td>.065</td>
<td>367</td>
</tr>
<tr>
<td><strong>Segment size (% of all households)</strong></td>
<td></td>
<td>19</td>
<td>0</td>
<td>5 8</td>
<td>3.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Price sensativity</strong></td>
<td></td>
<td>9 3</td>
<td>9 7</td>
<td>25.8</td>
<td>16 4</td>
</tr>
<tr>
<td>β</td>
<td></td>
<td>-.187</td>
<td>-.144</td>
<td>-.307</td>
<td>-.542</td>
</tr>
</tbody>
</table>

*For switching segments 1 through 4, purchase probabilities greater than .10 are underlined.

*Price coefficient statistically insignificant at the .05 level.

---

### Table 4

**HETEROGENEITY CORRECTION FACTORS**

<table>
<thead>
<tr>
<th>Segment</th>
<th>Mean correction</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>765</td>
<td>864</td>
<td>889</td>
<td>813</td>
<td>893</td>
</tr>
<tr>
<td></td>
<td>(.130)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>.808</td>
<td>832</td>
<td>840</td>
<td>766</td>
<td>872</td>
</tr>
<tr>
<td></td>
<td>(.052)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>798</td>
<td>766</td>
<td>736</td>
<td>753</td>
<td>936</td>
</tr>
<tr>
<td></td>
<td>(.072)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>755</td>
<td>715</td>
<td>712</td>
<td>661</td>
<td>766</td>
</tr>
<tr>
<td></td>
<td>(.056)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>819</td>
<td>803</td>
<td>808</td>
<td>821</td>
<td>789</td>
</tr>
<tr>
<td></td>
<td>(.031)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Standard deviations of correction factors are in parentheses. A correction factor equal to 1 implies segment homogeneity.
note that these predicted choice shares correspond well to the actual market shares in Table 1.

**Competitive Clout and Vulnerability**

The interpretation of the structure implied by the elasticity matrices is made easier if we define summary measures of brand competition. Suppose brand $i$ is of primary interest. Then, cross elasticities of the form $\eta_{ij}$ (the percentage change in $j$’s share with a 1% change in $i$’s price) report the ability of brand $i$ to take share away from competitors. By contrast, cross elasticities of the form $\eta_{ij}$ (the percentage change in $i$’s share with a 1% change in $j$’s price) report the vulnerability of brand $i$ to competitors.

Using these observations, we define for each brand $i$

$$\text{Competitive Clout} = \sum_{j \neq i} \eta_{ij}$$

and

$$\text{Vulnerability}, i = \sum_{j \neq i} \eta_{ji}$$

where the summation runs over all of brand $i$’s competitors. These measures are similar—though not identical—to Cooper’s (1988) notions of clout and receptivity.

**Table 5**

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>P</th>
<th>Share</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>Segment 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>-1.46</td>
<td>53</td>
<td>36</td>
<td>27</td>
<td>79.0</td>
</tr>
<tr>
<td>B</td>
<td>56.5</td>
<td>5.04</td>
<td>28</td>
<td>14</td>
<td>8.9</td>
</tr>
<tr>
<td>C</td>
<td>51.7</td>
<td>38</td>
<td>-4.60</td>
<td>19</td>
<td>6.9</td>
</tr>
<tr>
<td>P</td>
<td>56.8</td>
<td>28</td>
<td>28</td>
<td>-4.51</td>
<td>52</td>
</tr>
<tr>
<td>Segment 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>-4.01</td>
<td>2.76</td>
<td>34</td>
<td>17</td>
<td>21.9</td>
</tr>
<tr>
<td>B</td>
<td>1.13</td>
<td>-1.50</td>
<td>39</td>
<td>15</td>
<td>64.6</td>
</tr>
<tr>
<td>C</td>
<td>1.03</td>
<td>2.84</td>
<td>-3.64</td>
<td>13</td>
<td>9.2</td>
</tr>
<tr>
<td>P</td>
<td>1.18</td>
<td>2.50</td>
<td>31</td>
<td>-3.33</td>
<td>43</td>
</tr>
<tr>
<td>Segment 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>-8.50</td>
<td>2.06</td>
<td>4.06</td>
<td>57</td>
<td>15.6</td>
</tr>
<tr>
<td>B</td>
<td>1.51</td>
<td>-6.69</td>
<td>4.85</td>
<td>42</td>
<td>25.9</td>
</tr>
<tr>
<td>C</td>
<td>1.55</td>
<td>2.53</td>
<td>-4.11</td>
<td>47</td>
<td>52.0</td>
</tr>
<tr>
<td>P</td>
<td>1.93</td>
<td>1.94</td>
<td>4.10</td>
<td>-8.89</td>
<td>6.5</td>
</tr>
<tr>
<td>Segment 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>-15.07</td>
<td>3.25</td>
<td>3.64</td>
<td>4.72</td>
<td>9.5</td>
</tr>
<tr>
<td>B</td>
<td>1.57</td>
<td>-11.31</td>
<td>4.79</td>
<td>4.39</td>
<td>23.8</td>
</tr>
<tr>
<td>C</td>
<td>1.46</td>
<td>3.97</td>
<td>-10.19</td>
<td>4.89</td>
<td>30.1</td>
</tr>
<tr>
<td>P</td>
<td>1.69</td>
<td>3.25</td>
<td>4.38</td>
<td>-8.03</td>
<td>36.7</td>
</tr>
<tr>
<td>Total market</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>-2.17</td>
<td>65</td>
<td>72</td>
<td>34</td>
<td>35.5</td>
</tr>
<tr>
<td>B</td>
<td>1.08</td>
<td>-3.95</td>
<td>2.07</td>
<td>81</td>
<td>25.9</td>
</tr>
<tr>
<td>C</td>
<td>1.30</td>
<td>2.27</td>
<td>-4.51</td>
<td>1.24</td>
<td>24.8</td>
</tr>
<tr>
<td>P</td>
<td>1.20</td>
<td>1.73</td>
<td>2.42</td>
<td>-4.59</td>
<td>13.8</td>
</tr>
</tbody>
</table>

*Elasticities should be interpreted as the percentage change in the row brand share corresponding to a 1% change in the column brand price. Segments not shown here (switching segment 5 and all loyals) are assumed to have price elasticities equal to zero.

**Table 6**

<table>
<thead>
<tr>
<th>Competitive clout</th>
<th>Vulnerability</th>
<th>Choice share</th>
<th>Brand elasticity</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>Segment 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>90.9</td>
<td>.5</td>
<td>79.0</td>
<td>-1.46</td>
</tr>
<tr>
<td>B</td>
<td>5</td>
<td>32.0</td>
<td>8.9</td>
<td>-5.04</td>
</tr>
<tr>
<td>C</td>
<td>3</td>
<td>26.9</td>
<td>6.9</td>
<td>-4.60</td>
</tr>
<tr>
<td>P</td>
<td>1</td>
<td>32.4</td>
<td>5.2</td>
<td>-4.51</td>
</tr>
<tr>
<td>Segment 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>3.7</td>
<td>7.8</td>
<td>21.9</td>
<td>-4.01</td>
</tr>
<tr>
<td>B</td>
<td>21.9</td>
<td>1.5</td>
<td>64.6</td>
<td>-1.50</td>
</tr>
<tr>
<td>C</td>
<td>4</td>
<td>9.1</td>
<td>9.2</td>
<td>-3.64</td>
</tr>
<tr>
<td>P</td>
<td>1</td>
<td>7.7</td>
<td>4.3</td>
<td>-3.33</td>
</tr>
<tr>
<td>Segment 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>8.4</td>
<td>21.1</td>
<td>15.6</td>
<td>-8.50</td>
</tr>
<tr>
<td>B</td>
<td>14.4</td>
<td>26.0</td>
<td>25.9</td>
<td>-6.69</td>
</tr>
<tr>
<td>C</td>
<td>56.8</td>
<td>9</td>
<td>52.0</td>
<td>-4.11</td>
</tr>
<tr>
<td>P</td>
<td>7</td>
<td>24.3</td>
<td>6.5</td>
<td>-6.89</td>
</tr>
<tr>
<td>Segment 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>7.5</td>
<td>46.1</td>
<td>9.5</td>
<td>-15.07</td>
</tr>
<tr>
<td>B</td>
<td>36.9</td>
<td>44.7</td>
<td>23.8</td>
<td>-11.31</td>
</tr>
<tr>
<td>C</td>
<td>55.4</td>
<td>41.8</td>
<td>30.1</td>
<td>-10.19</td>
</tr>
<tr>
<td>P</td>
<td>65.5</td>
<td>32.6</td>
<td>36.7</td>
<td>-8.03</td>
</tr>
<tr>
<td>Total market</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>4.3</td>
<td>1.1</td>
<td>35.5</td>
<td>-2.17</td>
</tr>
<tr>
<td>B</td>
<td>8.6</td>
<td>6.1</td>
<td>25.9</td>
<td>-3.95</td>
</tr>
<tr>
<td>C</td>
<td>10.7</td>
<td>8.4</td>
<td>24.8</td>
<td>-4.51</td>
</tr>
<tr>
<td>P</td>
<td>2.3</td>
<td>10.3</td>
<td>13.8</td>
<td>-4.59</td>
</tr>
</tbody>
</table>

*Segments not shown here (switching segment 5 and all loyals) are assumed to have competitive clout and vulnerability measures equal to zero. Brand elasticity is the predicted own elasticity taken from Table 5.

The price changes of a brand with considerable competitive clout have a major impact on the shares of competitors. Conversely, a brand with a high vulnerability score will face relatively large changes in share in response to price changes of competitors. Thus, the strongest brands in the market would be expected to have high competitive clout and low vulnerability.

Competitive clout and vulnerability statistics for our data are reported in Table 6. At the segment level, there is a strong positive association between a brand’s competitive clout and its choice share, as well as a strong negative association between a brand’s vulnerability and its choice share. Evidently, competitive clout and vulnerability are complementary concepts, both driven by the brand’s share within the segment.

8 Using our definitions, we can write Cooper’s (1988) measures of clout and receptivity as

$$\text{Clout} = \text{Competitive Clout} + \eta_i^c$$

and

$$\text{Receptivity} = \text{Vulnerability} + \eta_i^c$$

Thus, in contrast to Cooper (1988), we ignore the own-price elasticities in defining a brand’s competitive position.
The link between preference segmentation and elasticity structure can be seen by plotting competitive clout versus vulnerability. In Figure 1, we represent each brand by a circle whose area is proportional to the average choice share within the segment. The negative association between competitive clout and vulnerability and the underlying causal connection to choice share can be seen easily. In fact, even if choice shares were unknown, it would be possible to infer the relative brand preferences within each segment by observing the competitive clout of each brand. Generally, segments with higher own-price elasticities also have larger measures of vulnerability (Table 6). Thus, we can use Figure 1 to conclude that segment 4 has the highest price sensitivity of all segments.9

Market-Level Competitive Clout and Vulnerability

The market-level elasticity matrix of Table 5 represents an average of the elasticity matrices of all seg-

---

9This relationship between overall price sensitivity and vulnerability must be interpreted cautiously because brands with low share also have high vulnerability. Notice that segment 1 has high vulnerability scores for small-share brands despite the fact that Table 3 shows its overall price sensitivity to be relatively low.
ments—including the brand-loyal segments. It is not, however, a simple weighted average. As equations 11 and 12 make clear, different weights are used for each row of the final matrix. Thus, relationships seen in a segment-level matrix may not be reflected at the market level.

In Figure 2, we display the competitive clout—vulnerability plot for all households in the study. Intuitively, we should expect the intensity of competition between brands to be larger within particular segments than in the entire market. For example, Table 5 shows high cross elasticities between the low-priced national brand C and private labels P for segment 4, the price-sensitive private label segment. However, when we aggregate over all segments, the market-level cross elasticities between these two brands are substantially smaller. This moderation in the size of the typical cross elasticity explains why the competitive clout and vulnerability scores in Figure 2 are considerably smaller than the corresponding scores in Figure 1.

Taking brands B and C as reference points, we see that the two most interesting brands are national brand A (low competitive clout, low vulnerability) and the private labels P (low competitive clout, high vulnerability). Though the low competitive clout of A is surprising in view of its market share, this outcome is actually a reflection of the brand's unusually large loyal segment (19% of the market) and its dominant position in switching segment 1. In fact, if the brand were purchased only by one completely loyal segment, competitive clout and vulnerability would be zero. Clearly, A is in a strong market position. It has the lowest elasticity of all brands

(Table 6) and is not affected by changes in other brands' prices.

In contrast, the weakness of the private labels is seen as low competitive clout combined with high vulnerability. That is, private labels have little impact on national brands, but are strongly affected by national brand price changes. This pattern is an excellent example of the Blattberg-Wisniewski (1985) price tier theory. The appearance of P in this position on the competitive clout—vulnerability plot can be attributed to the high price elasticity of segment 4 coupled with the limited appeal of private labels for national brand buyers.

CONCLUSIONS

We propose a new probabilistic choice model that uses observed purchase histories to classify a household's brand utilities (and purchase probabilities) into a small number of preference segments. In developing the model, we start with the choice process at the household level and make assumptions about the homogeneity of households within each segment. Thus, our model is compositional in the sense that it attempts to aggregate households into homogeneous preference segments. This approach stands in contrast to recently developed decompositional models (Cooper 1988; Grover and Srinivasan 1987; Shugan 1987), which attempt to identify market structure on the basis of aggregate data (market shares or brand switching).

Because our model estimates both intrinsic brand utilities and sensitivity to price changes, we make full use of the rich data available in current scanner panels and are able to develop a representation of market structure in terms of both brand preferences and price elasticities. For example, we show in our empirical work that the price-sensitive switching segments in one market can be structured in terms of price tiers. We also demonstrate how high intrinsic utility for one brand (A) can lead to a strong competitive advantage in terms of price elasticities. In general, our model is capable of monitoring the structure of a market and providing diagnostic information for the marketing manager.

In our study, we chose to concentrate on price as the most important destabilizing element of the marketing mix. Clearly, other variables could be added to the model to assess their impact on the preference structure. For example, by adding information about advertising, we could examine the short-run influence on brand choice and the long-term influence on the composition of the market's preference segments. The model's link between preferences and price sensitivity also provides a potential vehicle for commenting on the classical controversy over the impact of advertising on price elasticity (see, e.g., Farris and Albion 1980).

Two extensions of the model await further research. First, we could express a segment's intrinsic preference for brand j(ωi) as a function of segment-specific tastes and the brand's location in an estimated multiattribute space. The result would be a product map for the market that could be linked to both segment-level choice probabilities and reactions to price changes.

10In these calculations, we assume that brand loyals and switching segment 5 have elasticities equal to zero.
Second, as currently formulated, the model considers only the consumer's choice decision. Though this restriction does not affect the model's ability to develop the preference structure, it does limit the model to the prediction of choice share elasticities. Recent work on discrete-continuous choice (Krishnamurthi and Raj 1988) and on the determinants of sales elasticities (Russell and Bolton 1988) may enable the model to be reformulated to predict price sales elasticities. These issues are currently being investigated.

APPENDIX A

ESTIMATION OF THE MODEL

Using equation 6, we can write the log-likelihood for an observed set of K consumers as

\[
LL = \sum_i \log \left[ \sum_j \prod \exp(\lambda_i) L(H_{ij} | \theta) \right] - K \log \left[ \sum_i \exp(\lambda_i) \right]
\]

We obtained maximum likelihood estimates of \( \lambda_i, \beta_i, \) and \( \lambda \), using a modified Newton gradient search (Dennis and Schnabel 1983, Appendix A) to find the maximum of A1. The gradients required for this search are

\[
\frac{dLL}{d\lambda_i} = \sum_j \left\{ \left( \frac{h_j}{\sum_j h_j} \right) - \exp(\lambda_i) / \sum \exp(\lambda_j) \right\}
\]

\[
\frac{dLL}{d\beta_i} = \sum_j \left\{ \left( \frac{\exp(\lambda_i)}{\sum_j h_j} \right) \sum_i a_{ij} \right\}
\]

\[
\frac{dLL}{da_{ij}} = \sum_i \left\{ \left( \frac{\exp(\lambda_i)}{\sum_j h_j} \right) \sum_j b_{ij} \right\}
\]

where,

\[\begin{align*}
    h_j &= \exp(\lambda_i) L(H_{ij} | \theta), \\
    a_{ij} &= w_{ij}(X_i, \theta) / \sum_j w_{ij} q_i, \\
    b_{ij} &= w_{ij}(X_i, \theta) / \sum_j w_{ij}, \\
    w_{ij} &= \exp(u_{ij} + \beta_i X_{ij}), \\
    W_i &= \sum_j w_{ij}, \\
    Y_i &= \sum_j X_{ij} w_{ij}, \\
    q_i &= \prod_j w_{ij} / W_i, \text{ and} \\
    z_i(t) &= 1 - w_{ij} \text{ if brand } j \text{ is chosen at time } t
\end{align*}\]

Using the asymptotic properties of maximum likelihood estimates (see, e.g., Bickel and Doksum 1977, p 132), we can infer standard errors from the matrix of second derivatives evaluated at the point corresponding to the maximum of A1. In this study, these second derivatives were computed numerically.

APPENDIX B

CORRECTION FOR HETEROGENEITY

Suppose the market consists of \( J = 1, 2, \ldots, J \) brands. Define \( S_{ij}^{(t)} \) as the share of choices for brand \( j \) by consumer \( k \) and let \( \beta_i \) be this consumer's price sensitivity parameter. Assume that all \( k = 1, 2, \ldots, K \) consumers are members of segment \( i \) Then the choice share cross elasticities within segment \( i \) can be computed by extending the weighting rationale of equation 11 to aggregation within segments.

\[
\begin{align*}
    \tau_i^* &= \sum_k \left[ \frac{S_{ij}^{(t)}}{\sum_{k} S_{ij}^{(t)}} \right] (-\beta_i) \frac{S_{ij}^{(t)} X_{ij}}{\sum_i S_{ij}^{(t)}} \\
    &= \sum_k (-\beta_i) \frac{S_{ij}^{(t)} X_{ij}}{\sum_i S_{ij}^{(t)}} \\
    &= -E(\beta_i S_{ij}^{(t)} X_{ij}) / E(S_{ij}^{(t)}) \\
    &= -E(\beta_i S_{ij}^{(t)} X_{ij}) / S_{ij}
\end{align*}
\]

where \( E(\cdot) \) denotes expectation over all consumers in \( i \) and \( S_{ij} \) is the average choice probability for brand \( j \).

To evaluate this expression further, we assume that, conditional on the prices \( X_{ij} \), the distributions of \( \beta_i \) and \( S_{ij}^{(t)} \) are independent. That is, we assume that the knowledge of a consumer's price sensitivity \( \beta_i \) gives no information about his or her average preferences \( S_{ij}^{(t)} \). Though this assumption is unreasonable across an entire market, it is apt to be a good approximation within a relatively homogeneous segment. It will hold if all consumers in segment \( i \) are assumed to have the same price sensitivity parameter.

Using this assumption and letting \( E(\beta_i) = \beta_i \), we can express B1 as

\[
\begin{align*}
    \tau_i^* &= -\beta_i E(S_{ij}^{(t)} X_{ij}) / S_{ij} \\
    &= [E(S_{ij}^{(t)} X_{ij}) / S_{ij}] [1 - \beta_i X_{ij}] \\
    &= [E(S_{ij}^{(t)} X_{ij}) / S_{ij}] \tau_i^*
\end{align*}
\]

where \( \tau_i^* \) is the cross elasticity evaluated at the segment-level averages for \( \beta_i \) and \( S_{ij}^{(t)} \). This is the result stated in the text. The argument for the own-price elasticity is based on the relation

\[
\tau_i^* = \sum_k \left[ \frac{S_{ij}^{(t)}}{\sum_k S_{ij}^{(t)}} \right] [B_d (1 - S_{ij}^{(t)}) X_{ij}]
\]

Following the same process outlined above, we obtain

\[
\tau_i^* = [S_{ij} - E(S_{ij}^{(t)} X_{ij}) / (S_{ij} - S_{ij}^{(t)})] \tau_i^*
\]

REFERENCES


Recognizing Interdependence and Heterogeneous Preferences," *Econometrica*, 46 (March), 403–26


Reprint No. JMR264100