

Determining Segmentation in Sales Response across Consumer Purchase Behaviors



Randolph E. Bucklin; Sunil Gupta; S. Siddarth

Journal of Marketing Research, Vol. 35, No. 2. (May, 1998), pp. 189-197.

Stable URL:

<http://links.jstor.org/sici?sici=0022-2437%28199805%2935%3A2%3C189%3ADSISRA%3E2.0.CO%3B2-K>

Journal of Marketing Research is currently published by American Marketing Association.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/ama.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.

RANDOLPH E. BUCKLIN, SUNIL GUPTA, and S. SIDDARTH*

The authors develop a joint estimation approach to segment households on the basis of their response to price and promotion in brand choice, purchase incidence, and purchase quantity decisions. The authors model brand choice (what to buy) by multinomial logit, incidence (whether to buy) by nested logit, and quantity (how much to buy) by poisson regression. Response segments are determined probabilistically using a latent mixture model. The approach simultaneously calibrates sales response on two dimensions: across segments and the three purchase behaviors. The procedure permits market-level sales elasticities to be decomposed by segment and purchase behavior (i.e., choice, incidence, and quantity). The authors apply the approach to scanner panel data for the yogurt category and find substantial differences across segments in the relative impact of the choice, incidence, and quantity decisions on overall sales response to price.

Determining Segmentation in Sales Response Across Consumer Purchase Behaviors

When a consumer walks into a grocery store, he or she implicitly has to make three key decisions about a product category—whether to buy in the category; if so, which brand to buy; and what quantity to purchase. All three decisions are influenced by a consumer's own characteristics (e.g., brand loyalty, usage rate) and by the marketing environment (e.g., the price and promotion activity of the various brands). Marketing mix variables can affect these three decisions to differing degrees. For example, price might have a substantial influence on a consumer's brand choice decision but might not affect category purchase or quantity decisions. Moreover, these effects are likely to vary across shoppers. For example, a price reduction might induce a segment of consumers to switch brands with no effect on incidence and quantity while encouraging another segment of brand-loyal consumers to buy early or stockpile the product.

To decompose purchase behavior into the choice, incidence, and quantity effects or to segment purchase behavior

across households, analysis must be conducted at the disaggregate level (e.g., with the use of single-source scanner panel data). Such analysis—which is not possible with store-level data—has the capability to trace the sources of sales response to brand switching (choice), purchase acceleration (incidence), or stockpiling (quantity) and to group households according to their relative sensitivity to marketing activity. Although previous research has developed several approaches that address parts of this two-way decomposition, there is no approach that provides a fully integrated solution to this problem.

Our goal is to develop and illustrate an approach that provides a two-dimensional decomposition of sales response: by behavior (i.e., choice, incidence, and quantity) and segment. Our objective is to help researchers and practitioners unlock more of the diagnostic information available from the analysis of scanner panel data. Specifically, our modeling approach (1) captures consumers' decisions for brand choice, purchase incidence, and purchase quantity; (2) determines latent consumer segments on the basis of consumers' response to marketing activities across all three purchase decisions; and (3) jointly estimates models that describe the three behaviors and the latent response segments.

In our approach, brand choice is represented by multinomial logit (e.g., Guadagni and Little 1983), category purchase incidence is captured with nested logit (Bucklin and Lattin 1991), and purchase quantity is modeled with poisson regression (e.g., Maddala 1983). Estimating these three

*Randolph E. Bucklin is an associate professor, Anderson Graduate School of Management, University of California, Los Angeles (e-mail: rbucklin@anderson.ucla.edu). Sunil Gupta is a professor, Graduate School of Business, Columbia University (e-mail: sg37@columbia.edu). S. Siddarth is a visiting professor, Anderson Graduate School of Management, University of California, Los Angeles (e-mail: s.siddarth@anderson.ucla.edu). The authors thank the ACNielsen company for the data used in this study and Louis P. Bucklin, Jim Lattin, Seenu Srinivasan, three anonymous *JMR* reviewers, and the editor for helpful comments.

components together enables investigators to decompose the impact of marketing variables on choice, incidence, and purchase quantity. Nevertheless, this type of decomposition remains tied to an aggregate-level view because all panelists continue to share the same response parameters for marketing variables.

To relax this constraint, the second dimension of our decomposition incorporates differences in response behavior across households. We use a probabilistic mixture model that simultaneously estimates the response parameters for choice, incidence, and quantity by latent segment. The mixture model approach determines response segments post hoc, on the basis of the revealed purchase behavior of the panelists in response to observed variation in marketing activity (e.g., Kamakura and Russell 1989).

Some previous studies on scanner panel data have considered all three purchase decisions, but none of the methods has incorporated segmentation in response to marketing activity. For example, Gupta (1988) considers all three purchase behaviors but does not consider segmentation. Chiang (1991) and Chintagunta (1993) build models that handle choice, incidence, and quantity but do not address response segmentation. In addition, both Chiang's and Chintagunta's approaches impose a modeling restriction on the price coefficient, such that the price elasticities for the purchase quantity decision are constrained to be equal to (or within estimation error of) -1.0 . Our use of the poisson regression model for quantity enables us to avoid this problem.

Our approach explicitly considers response segmentation through a latent class model (i.e., post hoc instead of a priori) that simultaneously incorporates all three purchase behaviors. Although several latent class models have appeared in the marketing literature, none has incorporated choice, incidence, and quantity at the same time. For example, Kamakura and Russell (1989) and Jain, Vilcassim, and Chintagunta (1994) consider only brand choice, and Bucklin and Gupta (1992) consider brand choice and purchase incidence. Bucklin and Gupta's results also are based on a sequential estimation (choice, then incidence), which can lead to slightly less efficient estimates than a simultaneous estimation (Ben-Akiva and Lerman 1985). In this study, we base the segmentation results on a joint and simultaneous estima-

tion of choice, incidence, and quantity parameters. Table 1 presents a simple taxonomy of some of the key methodological research leading up to the model proposed here.

A key feature of our procedure is the joint estimation of segmentation and response behaviors. Alternatively, latent segments for each behavior can be estimated separately, which yields S_c choice segments, S_i incidence segments, and S_q quantity segments. With this approach, panelists would be classified into one of $S_c \cdot S_i \cdot S_q = S_{sep}$ segments, which represent overall response. One drawback of a separate procedure is the potentially large number of response segments produced (e.g., if each response behavior was described with three segments, a household could belong to one of $3 \cdot 3 \cdot 3 = 27$ different groups). Another drawback of a separated procedure is the assumption that the segmentation for each response behavior is independent of the other two.

We apply our segmentation approach to scanner panel data for the yogurt category. We estimate both the joint and separate segmentation schemes and compare the results. We also use the joint approach to provide a decomposition of price elasticities by segment and response behavior, generalizing the decomposition first proposed by Gupta (1988). We show that the relative contribution of choice, incidence, and quantity to total price elasticity differs substantially across the segments.

MODELING APPROACH

The objective of our model is to capture the impact of marketing variables on consumers' decisions for choice, incidence, and quantity and to represent the cross-sectional heterogeneity in response to those variables parsimoniously. The approach enables us to capture a wide variety of different purchase patterns. For example, we might obtain a segment of consumers that is price sensitive in choice but insensitive in incidence and quantity. These consumers switch brands because of temporary price reductions, but they do not accelerate their purchases or buy more than their usual quantity. A second segment of consumers might be brand loyal, shop at regular time intervals, and take advantage of price discounts on a favorite brand by stockpiling. Our model would classify these consumers as price insensitive in choice and incidence and price sensitive in quantity.

Table 1
RESEARCH BACKGROUND

<i>Behavior</i>	<i>Unsegmented</i>	<i>A Priori Segmentation</i>	<i>Post Hoc Segmentation</i>
<i>Choice</i>	Guadagni and Little (1983)	Gensch (1985)	Kamakura and Russell (1989) Jain, Vilcassim, and Chintagunta (1994)
<i>Incidence</i>	Guadagni and Little (in press)	Neslin, Henderson, and Quelch (1985) Bucklin and Lattin (1991)	Bucklin and Gupta (1992)
<i>Quantity</i>	Neslin, Henderson, and Quelch (1985) Krishnamurthi and Raj (1988)	Krishnamurthi and Raj (1991)	
<i>All Three</i>	Gupta (1988) Chiang (1991) Chintagunta (1993)		This Study

Consumers who buy early, switch brands, and stockpile the promoted brand are labeled as price sensitive in all three decisions. A key feature of our approach is that we do not prespecify these segments but use the actual purchase data of households to classify them on the basis of their response behavior.

We begin by describing the components of a single-segment model. We then describe the multiple-segment model and estimation approach.

Single-Segment Model

We conceptualize the consumer purchase decision as follows. Conditional on a store visit, a shopper decides whether to make a purchase from a product category. If a category purchase is to be made, the consumer then decides which brand to buy. Given brand choice, the consumer decides how many units of that brand to purchase. We translate this into a model by decomposing the unconditional probability that a household buys q_i units of brand i on a shopping trip as the product of the purchase incidence probability (modeled with nested logit), the brand choice probability given incidence (modeled with multinomial logit), and the probability of buying q_i units given incidence and the choice of brand i (modeled with poisson regression). Our model is consistent with an incidence \rightarrow choice \rightarrow quantity decision sequence, but it does not assume that consumer decisions necessarily proceed in that order.

Brand choice model. The probability that household h buys brand i on a store visit at time t , given a decision to buy in the product category (i.e., conditional on purchase incidence), is given by the multinomial logit expression (e.g., Guadagni and Little 1983)

$$(1) \quad P_t^h(i|inc) = \frac{\exp(u_i + \beta X_{it}^h)}{\sum_k \exp(u_k + \beta X_{kt}^h)}$$

Given a vector of marketing and household-specific variables X_{it}^h , the model specifies the estimation of brand-specific intercepts u_i and a vector of response coefficients β . Household-specific variables, such as brand loyalty, are included to capture heterogeneity in consumer choice probabilities.

Purchase incidence model. On each shopping trip, a consumer decides whether to buy from a product category. This buy/no-buy outcome can be modeled with binary nested logit, in which the probability of purchase incidence for household h on a store visit at time t is given by (e.g., Bucklin and Gupta 1992)

$$(2) \quad P_t^h(inc) = \frac{\exp(\gamma_0 + \gamma Y_t^h)}{1 + \exp(\gamma_0 + \gamma Y_t^h)}$$

Here, γ_0 is the intercept and γ is the vector of response coefficients for the explanatory variables Y_t^h . In the nested logit approach, the brand-choice decision is nested in the purchase-incidence decision. This is modeled by including category value as an explanatory variable in the vector Y_t^h . Category value, CV_t^h , is the maximum expected utility available to household h from buying a brand in the cate-

gory at time t and is given by the log of the denominator of the brand-choice probability (Ben-Akiva and Lerman 1985),

$$(3) \quad CV_t^h = \ln \sum_k \exp(u_k + \beta X_{kt}^h)$$

Nested logit is an appealing model for the category purchase decision for several reasons. It can be derived from a theory of stochastic utility maximization (Ben-Akiva and Lerman 1985). It also has the property that a household's incidence decision is influenced more by the marketing activity of its preferred brands than of its nonpreferred brands (e.g., Bucklin and Gupta 1992). Estimation is also straightforward and can be performed sequentially or simultaneously with the brand-choice model.

Purchase quantity model. Given purchase incidence and choice of brand i , the probability that household h buys $q_{it}^h = 1, 2, \dots, n$ units of brand i on a store visit at time t is captured by a poisson model with truncation of the zero outcome. This can be written as

$$(4) \quad P(Q_{it}^h = q_{it}^h | Q_{it}^h > 0) = \frac{\exp(-\lambda_{it}^h)(\lambda_{it}^h)^{q_{it}^h}}{[1 - \exp(-\lambda_{it}^h)]q_{it}^h!}$$

where λ_{it}^h is the purchase rate of household h for brand i at time t . We expect that the purchase rate λ_{it}^h will differ across households, brands, and in response to marketing activity. To incorporate these effects, we model λ_{it}^h as a function of both household-specific variables and marketing activity, so that

$$(5) \quad \lambda_{it}^h = \exp(\theta_0 + \theta Z_{it}^h)$$

where Z_{it}^h is a vector of household-specific and marketing variables and θ is a vector of parameters to be estimated. We specify λ_{it}^h as an exponential function of the explanatory variables to ensure that the purchase rate will be nonnegative. Unlike some previous research in modeling choice and quantity (e.g., Krishnamurthi and Raj 1988), we do not model the correlation in errors between choice and quantity.

We believe the poisson approach has several appealing properties. In many product categories, the number of units purchased is a more natural unit of analysis than ounces (e.g., frozen entrees, bathroom rolls, soups, air fresheners, yogurt). Modeling quantity in ounces also typically involves a regression approach in which the dependent variable is treated as continuous (Krishnamurthi and Raj 1988; Neslin, Henderson, and Quelch 1985). Because quantity purchased is often discrete, treating discrete outcomes as continuous can result in biased estimates (McKelvey and Zavoina 1975). Brand choice models also sometimes use the brand size (not just the brand) as the dependent variable, because different sizes of the same brand frequently are promoted differently (e.g., Guadagni and Little 1983). A quantity model that predicts units purchased therefore should be more compatible with choice models that predict brand sizes than with a model that predicts a continuous outcome, such as number of ounces.

Table 2
BRANDS, MARKET SHARES, AND AVERAGE PRICES

Brand	Market Share	Average Price	Average Price/ounce
Yoplait	.19	58.1 cents	9.7 cents
Dannon	.11	67.1 cents	8.4 cents
Weight Watcher's	.07	60.8 cents	7.6 cents
Nordica	.22	39.1 cents	6.5 cents
W.B.B.	.09	41.1 cents	5.1 cents
QC	.07	35.8 cents	4.5 cents
Private Label	.14	36.7 cents	4.6 cents

With the specification of the choice, incidence, and quantity components, we now can write the unconditional probability of household h buying q_{it}^h units of brand i on a store visit at time t as

$$(6) \quad P(Q_{it}^h = q_{it}^h) = P(Q_{it}^h = q_{it}^h | Q_{it}^h > 0) P_i^h(i|inc) P_i^h(inc).$$

Multiple-Segment Model

In the multiple-segment model, the unconditional probability of buying q_{it}^h units is given by the weighted sum of segment-level probabilities for quantity, choice, and incidence. That is,

$$(7) \quad P(Q_{it}^h = q_{it}^h) = \sum_{s=1}^S \pi_s P_s(Q_{it}^h = q_{it}^h | Q_{it}^h > 0) P_{st}^h(i|inc) P_{st}^h(inc),$$

where

$$\begin{aligned} \pi_s &= \text{share of segment } s, 0 < \pi_s \leq 1, \\ &\quad \sum \pi_s = 1; \\ S &= \text{number of segments (unknown);} \\ P_s(Q_{it}^h = q_{it}^h | Q_{it}^h > 0) &= \text{quantity probability for } h \text{ in segment } s; \\ P_{st}^h(i|inc) &= \text{choice probability for } h \text{ in segment } s; \\ P_{st}^h(inc) &= \text{incidence probability for } h \text{ in segment } s. \end{aligned}$$

The segment-level choice, incidence, and quantity probabilities are modeled similarly to Equations 1–4, and the purchase rate is modeled as in Equation 5, with the exception that model parameters are now segment specific.¹

Estimation

The log-likelihood (LL) function for the jointly estimated multisegment model is given by

$$(8) \quad LL = \sum_h \ln \left(\sum_s \pi_s \prod_i \left\{ P_{st}^h(inc) \delta_{it}^h [1 - P_{st}^h(inc)]^{1 - \delta_{it}^h} \prod_i \left[P_{st}^h(i|inc) P_s(Q_{it}^h = q_{it}^h | Q_{it}^h > 0) \right]^{\delta_{it}^h} \right\} \right),$$

where δ_{it}^h if household h buys brand i on a store visit at time t , and 0 otherwise, and $\delta_{it}^h = \sum_i \delta_{it}^h$.

The parameters of the multisegment model are estimated, conditional on a prespecified number of latent segments. We proceed by first calibrating a single-segment model, then a two-segment model, a three-segment model, and so forth, until the parameters needed to describe an additional segment no longer contribute significantly to model fit. As a model selection criterion, we employ the Bayesian Information Criterion (BIC), given by $BIC = LL - (k/2)\ln(n)$, where k is the number of parameters and n is the sample size.

EMPIRICAL APPLICATION

In this section, we present an application of our model to households that purchase in the yogurt category. We begin by describing the data set. We then specify the variables used in our model. Next, we discuss model calibration and segment selection on the basis of our joint approach, as well as the alternative separate estimation approach. We then present parameter estimates for the joint model.

Data

The data are from ACNielsen scanner panel records for households in Sioux Falls, S. Dak., during the period 1986–1988. We use the last 52 weeks of the period for model calibration and the preceding 61 weeks to initialize some model variables. For our estimation sample, we randomly selected 300 yogurt-purchasing households. During the estimation year, these households made a total of 30,791 store trips at 13 stores in Sioux Falls. Some 3485 store trips resulted in yogurt category purchases, and 82% of purchases were for multiple units.

To keep the model estimation tractable, we limited our study to the top seven selling yogurt brands in the market. Together, these accounted for 84% of category sales in dollars and 89% in units. These brands, their market shares in units, and average prices are given in Table 2. The private-label brand was offered in one chain of stores only. Each brand was offered in one size, with all brands available in eight-ounce containers, except for Yoplait and Nordica, which were available in six-ounce containers.

Model Variables

We now turn to the specification of variables for the three components of the model.

Choice. The deterministic component of utility for brand i at time t , given that household h is a member of segment s , is specified as

$$(9) \quad U_{ist}^h = u_{is} + \beta_{s1} BL_i^h + \beta_{s2} LBP_{it}^h + \beta_{s3} PRICE_{it} + \beta_{s4} PROMO_{it}$$

¹The segment size parameter, π_s , is estimated by reparameterizing it as $\exp(\phi_s) / \sum_s \exp(\phi_s)$.

Brand loyalty (BL_i^h) is the within-household market share of each brand during the 61-week initialization period. Our measure for brand loyalty varies across households but remains constant through the calibration period and is designed to capture purely cross-sectional heterogeneity. Last brand purchased (LBP_{it}^h) measures the time-varying component of a household's brand preference and is designed to capture such effects as purchase event feedback.²

We model the effects of two marketing variables: price and the presence or absence of nonprice promotional activity. Our measure for the price of the brand chosen, $PRICE_{it}$, is the actual shelf price (including temporary discounts) in cents per ounce. For reasons of parsimony, we capture the promotional status of a brand i , $PROMO_{it}$, with an indicator variable set equal to one if the brand is featured or displayed at occasion t and set equal to zero otherwise.³

Incidence. The utility derived by household h for making a product purchase during a store visit at time t , given that the household belongs to segment s , is expressed as

$$(10) \quad V_{st}^h = \gamma_{s0} + \gamma_{s1}CR^h + \gamma_{s2}INV_t^h + \gamma_{s3}CV_{st}^h$$

Consumption rate, CR^h , is a household's average weekly consumption of yogurt. This is computed as the total amount of yogurt purchased by a household h in the initialization period divided by the number of weeks in the period. It remains constant through the calibration period. Households with higher consumption rates should be more likely to purchase yogurt on a given store visit than households with lower consumption rates, and we expect a positive sign for γ_{s1} .

The inventory variable, INV_t^h , is designed to capture time-varying heterogeneity in incidence probabilities. Households with high at-home inventories of yogurt should be less likely to buy at their next store visit. To construct the inventory variable, we assume that households draw down their supply of yogurt linearly at their rates of consumption, CR^h . We initialize our inventory measure at zero at the start of the initialization period. We also mean-center INV_t^h by subtracting each household's average level of inventory during the calibration period. This makes the measure purely longitudinal, so that INV_t^h becomes a measure of relative inventory in a household. We expect a negative sign for γ_{s2} .

In the nested logit incidence model, the attractiveness of the product category (due to the price and promotion of the various brands) is represented by category value, CV_{st}^h (the log of the denominator of the brand-choice model). The coefficient for category value should be positive (reflecting the positive influence of, for example, a price promotion on the incidence decision) and lie between zero and one so that it is consistent with stochastic utility maximization theory (Ben-Akiva and Lerman 1985).

Quantity. The poisson rate parameter for brand i and household h belonging to segment s at time t , is

$$(11) \quad \lambda_{ist}^h = \exp(\theta_{is0} + \theta_{s1}PR^h + \theta_{s2}INV_t^h + \theta_{s3}BL_i^h + \theta_{s4}PRICE_{it} + \theta_{s5}PROMO_{it}).$$

Purchase rate, PR^h , is defined as the average quantity of yogurt purchased by household h , given that a yogurt purchase was made.⁴ It is computed as the total number of units purchased in the initialization period divided by the number of yogurt purchase occasions in the period. This variable is designed to capture heterogeneity in purchase propensity across households. A household with a high purchase rate is likely to buy more units on a given purchase occasion than a household with a low purchase rate. We expect θ_{s1} to be positive. Inventory, as was defined previously, also is included in the purchase-quantity model, because a large supply at home might make shoppers less likely to purchase additional units.⁵

The three remaining variables in Equation 11 are brand specific, pertaining to the particular brand the household selected at the purchase occasion. Brand loyalty, described previously, is included to capture the potential impact of brand preference on purchase quantity. We hypothesize that some households will tend to purchase more units of their favored brands and therefore anticipate a positive sign for θ_{s3} . The price and promotion variables capture marketing activity and are specified as in the choice model.

Calibration of Joint Model and Segment Selection

The choice of starting values is important in this application because of the large number of observations (30,791 store visits), large number of parameters, and three nonlinear model components. We therefore develop a simple multistep approach to obtain "good" starting values for the maximization of the joint likelihood function. We first estimate the parameters of the brand choice model alone. We then estimate the poisson regression quantity model, conditional on brand choice. Using the obtained choice and quantity parameters as starting values, we estimate the parameters of these two models jointly. Using the new choice parameters, we obtain category value (Equation 3) and estimate the incidence model. With starting values for each component, we then estimate all model parameters simultaneously.

The procedure for estimating a multisegment model is similar. Conditional on the number of segments, we begin by calibrating a multisegment choice model, then the multisegment quantity model, and finally the multisegment incidence model. When we have obtained starting values from our sequential approach, we estimate all parameters simultaneously by maximizing the likelihood in Equation 8. We

²The correlation between brand loyalty and last brand purchased is .37 ($p < .01$).

³The correlation between feature and display is .29 ($p < .01$). Coupon usage is insignificant in the yogurt category (5% of purchases) and is ignored.

⁴The correlation between purchase rate and consumption rate is .57 ($p < .01$).

⁵In the empirical application reported here, we find that the inventory coefficient in the purchase quantity model is consistently close to zero and statistically insignificant. We therefore drop it from this model. We have, however, found inventory to be significant in other applications of this purchase quantity model.

Table 3
SUMMARY OF CALIBRATED MODELS

	Number of Segments					
	1	2	3	4	5	6
<i>Joint Estimation</i>						
Number of parameters	25	38	51	64	77	90
Calibration LL	-19,950.3	-19,335.3	-19,087.6	-18,944.6	-18,851.7	-18,847.7
BIC	-20,079.5	-19,531.7	-19,351.1	-19,275.3	-19,249.6*	-19,312.8
<i>Choice Only</i>						
Number of parameters	10	15	20	25	30	—
Calibration LL	-3,913.4	-3,702.2	-3,554.8	-3,496.9	-3,484.9	—
BIC	-3,954.2	-3,763.4	-3,636.4	-3,598.9	-3,607.2	—
<i>Incidence Only</i>						
Number of parameters	4	9	14	19	24	—
Calibration LL	-10,103.3	-9,582.0	-9,361.7	-9,304.2	-9,302.3	—
BIC	-10,124.0	-9,628.5	-9,434.0	-9,402.4	-9,426.3	—
<i>Quantity Only</i>						
Number of parameters	11	16	21	26	—	—
Calibration LL	-5,936.9	-5,842.5	-5,791.0	-5,788.8	—	—
BIC	-5,981.8	-5,907.7	-5,876.6	-5,894.8	—	—

*Boldface indicates model selected.

repeat the procedure to obtain the parameters associated with each number of latent segments desired.

The top portion of Table 3 summarizes the estimation results for the joint, three-behavior model. The one-segment model has a total of 25 parameters (10 for choice, 4 for incidence, and 11 for quantity) and a fit of $-19,950.3$ log-likelihood points. The two-segment model adds 13 parameters (4 for choice, 4 for incidence, 4 for quantity, and 1 for segment size) and shows a substantial improvement in fit to $-19,335.3$ log-likelihood points.⁶ Results for the three-, four-, five-, and six-segment models also are reported in Table 3. The BIC is maximized at five segments, and we choose the five-segment solution as the basis for further analysis and sales response decomposition.

Comparison with Separate Estimation Approach

As we noted previously, an alternative approach to segmenting panelists estimates response segmentation separately for each behavior. In this method, estimation and segmentation are used for choice alone, incidence alone, and quantity alone. The lower portion of Table 3 reports the

results from estimating latent class models for choice, incidence, and quantity separately.⁷ Using the BIC, a separate estimation procedure would select four choice segments, four incidence segments, and three quantity segments.⁸

Because the approach assumes that the probabilistic components in all three models are independent, the resulting number of segments is $4 \cdot 4 \cdot 3 = 48$. Segment membership probabilities are given by the product of panelists' posterior membership probabilities corresponding to the 48 possible combinations of choice, incidence, or quantity segments. Most of the resulting segments are small; 30 of the 48 have posterior sizes less than or equal to 1%.

The fit of the separate approach can be assessed by adding the component likelihoods from the four-segment choice, four-segment incidence, and three-segment quantity models in Table 3. This gives a log-likelihood of $-18,592.1$ (versus $-18,851.7$ for the joint approach). A trade-off therefore emerges between the parsimony of the joint approach versus the fit of the separate approach. (The results of the separate approach do not reflect any correlation in response behaviors across segments because of the independence assumption.)

We further examine the trade-off between the two methods by studying several representations derived from the separate segmentation approach. We were interested in knowing how many of the 48 segments (proceeding from largest to smallest) would be needed to match the fit of the five-segment solution from the joint approach. We took the largest 5, 10, and 15 segments of the 48 and computed the

⁶To make the model estimation tractable, additional vectors of brand-specific constants were not added to the choice and quantity models for each additional segment. This reduces the number of additional parameters per segment by 13 (6 choice constants plus 7 quantity constants). However, for each multisegment model, the choice and quantity constants are reestimated; that is, they are not fixed at their values from the one-segment solution. We tested the impact of this restriction at two levels. First, we estimated a two-segment joint model, including vectors of brand constants for each segment. This model had $38 + 13 = 51$ parameters, a calibration log-likelihood of $-19,294.5$, and a BIC of $-19,568.5$, which indicated that the restricted model in Table 3 (BIC = $-19,531.7$) is preferred. Second, we looked at the brand-choice component alone, adding brand intercepts to each segment. A representative result is given by the four-segment solution in which the unrestricted model has $25 + 18 = 43$ parameters, a calibration log-likelihood of $-3,396.1$, and a BIC of $-3,571.4$. Thus, when considering choice alone, the unrestricted model is preferred.

⁷For the purchase-incidence model, the estimated parameters from the one-segment choice model were used to obtain category value. These are not included in the count of parameters in the incidence model because they are not reestimated.

⁸The choice segment sizes were .19, .10, .42, and .29; the incidence segment sizes were .23, .26, .44, and .07; and the quantity segment sizes were .15, .17, and .68.

posterior fit associated with those representations.⁹ The results are summarized in Table 4. An analysis of the 11 largest segments shows that a “break-even” in fit occurs between 10 and 11 separate-approach segments, versus the joint approach and its five-segment solution. We expect this trade-off to vary across product categories, favoring the joint approach when response segmentation is more correlated across behaviors and the separate approach when response behavior segmentation is largely independent across behaviors.

⁹The posterior likelihood is obtained by computing each household's posterior probabilities of segment membership and using those vectors in place of the priors, $\{\pi_i\}$. In the case of the separate approach, the posterior probabilities are computed using the parameter vectors that describe the 5, 10, 11, or 15 largest segments and are normalized to sum to one. Note also that the posterior likelihood for the five-segment joint approach (-18,517.2) will run higher than the corresponding prior likelihood in Table 3 (-18,851.7).

Table 4
COMPARISON OF POSTERIOR FIT FOR SEPARATE AND JOINT APPROACHES

Method	Number of Segments	Posterior Log-Likelihood
Separate	5 largest	-19,080.7
Separate	10 largest	-18,557.5
Joint	5	-18,517.2
Separate	11 largest	-18,439.2
Separate	15 largest	-18,293.7

Parameter Estimates from the Joint Approach

In Table 5, we present parameter estimates that are based on the joint model for both the single-segment solution (labeled aggregate), and the selected five-segment solution. All statistically significant parameters, with one exception, are in the hypothesized direction.¹⁰

Parameter estimates for the various segments provide a useful description of the response profile of the various segments.¹¹ Responsiveness to the marketing mix elements varies both across behaviors in a segment and over the different segments. For example, Segment 1 (7% of households, accounting for 17% of volume) is moderately price sensitive in choice but unaffected by promotion. In the quantity model, the promotion parameter for this segment also indicates a lack of promotion sensitivity.

A comparison of the segment-level results with the aggregate- (single-segment) level results shows several interesting differences. The aggregate results show a positive and significant effect of promotion on brand choice, as well as

¹⁰The consumption rate parameter in Segment 5 is incorrectly signed. This result might be due to chance (with 77 parameters, approximately four might be statistically significant at the .05 level, even though the true parameter is not different from zero). An alternative explanation is that consumption rate, computed from actual purchases made in the initialization period, is a poor measure of consumption rate for some households in the calibration period.

¹¹Although Segment 2 is small (3% of the sample or .03 · 300 = 9 panelists), the latent class procedure uses all the data when estimating this segment's parameters.

Table 5
PARAMETER ESTIMATES FROM JOINT ESTIMATION RESULTS FOR AGGREGATE AND FIVE-SEGMENT MODELS

Variables	Aggregate	Segment 1	Segment 2	Segment 3	Segment 4	Segment 5
Brand-Choice Model						
Brand loyalty	2.206 (28.19) ^a	4.742 (18.06)	5.046 (9.19)	1.764 (13.81)	2.503 (12.55)	.931 (4.54)
Last brand purchased	1.037 (23.65)	-.357 (-2.27)	.182 (.64)	.557 (6.79)	.674 (5.59)	2.625 (20.75)
Price	-.278 (-8.96)	-.244 (-5.69)	-.297 (-3.90)	-.441 (-13.81)	-.311 (-8.25)	-.154 (-3.80)
Promotion	1.536 (22.59)	.363 (1.47)	2.380 (5.74)	1.621 (15.95)	1.946 (13.40)	.684 (2.71)
Purchase-Incidence Model						
Consumption rate	1.988 (23.75)	.860 (4.97)	1.441 (1.74)	1.908 (11.88)	.824 (2.96)	-2.923 (-6.08)
Inventory	-.163 (-5.78)	-.144 (-2.45)	-2.725 (-5.04)	-.126 (-2.54)	-.583 (-8.13)	.049 (.77)
Category value	.503 (15.29)	.225 (2.98)	.211 (2.15)	.624 (11.07)	.543 (9.80)	.020 (.19)
Purchase-Quantity Model						
Purchase rate	.748 (32.48)	.770 (18.21)	1.925 (8.48)	.730 (16.91)	.912 (14.72)	.975 (8.98)
Brand loyalty	.250 (7.27)	.640 (6.68)	.093 (.25)	.214 (3.56)	.053 (.70)	.367 (4.42)
Price	-.064 (-4.64)	-.086 (-4.92)	-.333 (-5.42)	-.051 (-3.29)	-.066 (-3.98)	-.049 (-2.97)
Promotion	.147 (5.40)	-.112 (-1.44)	.287 (1.79)	.201 (5.51)	.167 (3.62)	.118 (1.27)
Segment Size						
Households	1.00	.07	.03	.35	.44	.11
Volume	1.00	.17	.03	.43	.20	.17

^aAsymptotic t-statistics in parentheses.

Table 6
AGGREGATE AND SEGMENT-LEVEL PRICE ELASTICITIES BY PURCHASE BEHAVIOR

<i>Behavior</i>	<i>Aggregate^a</i>	<i>Segment 1</i>	<i>Segment 2</i>	<i>Segment 3</i>	<i>Segment 4</i>	<i>Segment 5</i>
Choice	-1.08 (58.38) ^b	-.70 (47.55)	-.71 (38.15)	-1.47 (63.63)	-.94 (52.37)	-.42 (56.09)
Incidence	-.36 (19.46)	-.18 (11.90)	-.18 (9.69)	-.59 (25.36)	-.52 (28.89)	-.01 (1.69)
Quantity	-.41 (22.16)	-.60 (40.55)	-.96 (52.17)	-.25 (11.00)	-.34 (18.75)	-.31 (42.22)
Total	-1.85 (100.00)	-1.48 (100.00)	-1.85 (100.00)	-2.31 (100.00)	-1.80 (100.00)	-.74 (100.00)
Segment Size						
Households	1.00	.07	.03	.35	.44	.11
Volume	1.00	.17	.03	.43	.20	.17

^aAggregate elasticities are based on the parameters from the single-segment joint model.

^bPercentage of the total.

on purchase quantity. But segment-level results for choice suggest that 7% of households (accounting for 17% of volume) are not responsive to promotion. In quantity, 21% of households (accounting for 37% of volume) are not responsive to promotion. The aggregate model also finds a significant, positive impact of brand preference on quantity. But the segment-level results show that this is true for only 53% of the households (Segments 1, 3, and 5). These patterns of response heterogeneity illustrate the value of conducting a segment-by-behavior decomposition of sales response.¹²

PRICE ELASTICITY ANALYSIS

In this section, we analyze the sales response to price for our sample of households purchasing in the yogurt category. We use elasticity as our measure of sales response to price, computing elasticity in each of the five jointly estimated response segments. We then decompose the overall sales elasticity into the response due to choice, incidence, and quantity decisions.

Using the actual marketing environment during the calibration period, we compute the posterior probability of segment membership for each household. Using these segment membership probabilities and the segment-level parameters for the choice, incidence, and quantity models, we compute a "baseline" sales figure for each brand in each segment for the calibration period. We then proceed to compute arc elasticities by reducing the price of each brand by 1% and assessing the impact on sales.

To decompose the source of the overall elasticity into choice, incidence, and quantity, we proceed as follows: To examine the impact of choice alone, for example, we hold the incidence and quantity components at the baseline level and allow only the choice probability to change. Although the incidence and quantity components are fixed, they serve as weights for each household so that a heavy buyer will

have more influence on the average choice elasticity than a light buyer. We then simulate the new sales for each brand in each segment. Because we are using a one percentage point change in price, price elasticity is given by (new sales - baseline sales)/baseline sales. We use a similar procedure to compute the price elasticity for incidence and quantity. For incidence, we hold choice and quantity fixed, but include them as weights, whereas for quantity, we hold choice and incidence fixed but include them as weights. The overall price elasticity for a given segment can be decomposed into the sum of the choice, incidence, and quantity elasticities for that segment (e.g., Gupta 1988, Appendix B).

Table 6 presents both an aggregate-level decomposition for category-level price elasticity and a segment-by-segment decomposition. In the first column, we present the aggregate-level results based on the parameters from the single-segment solution. Note that this column decomposes response for all households into the choice (58%), incidence (20%), and quantity (22%) components. This result is analogous to the decompositions provided by Gupta (1988), who worked with data on grocery coffee, and by Chintagunta (1993), who also worked with yogurt data. We note that this 58/20/22 decomposition appears to differ by both product category and research method. Gupta (1988) reports an 84/14/2 decomposition for coffee; this suggests that marketing actions have a smaller impact on the quantity decision for coffee than for yogurt, perhaps due to package size conventions or perishability. Chintagunta's (1993) results imply a decomposition of approximately 40/15/45. Although his choice and incidence elasticities are close to those we report, our quantity elasticity (-.41) is substantially lower than his figure of approximately -1.0. As was noted previously, Chintagunta's approach imposes a modeling restriction on the price elasticity for purchase quantity.

The balance of Table 6 shows how our two-dimensional decomposition of sales response is reflected in price elasticities. The multisegment solution permits both the overall price elasticity and the relative impact of the purchase behavior components to vary across the segments. The results show that this intersegment variation is substantial. The impact of the choice decision ranges from 38% (Segment 2) to

¹²As a check on the stability of the joint model segmentation, we computed the posterior segment sizes for a holdout sample of 300 households using the parameters estimated on the calibration sample (Table 5). We found a close correspondence; the holdout segment sizes are equal to 7%, 2%, 32%, 47%, and 12%, respectively.

64% (Segment 3) of the total elasticity. The impact of incidence ranges from 10% (Segment 2) to 29% (Segment 4), whereas the impact of quantity ranges from 11% (Segment 3) to 52% (Segment 2).

Segments 3 and 4 are characterized by high price sensitivity in choice (52–64%), moderate sensitivity in incidence (25–29%), and low sensitivity in quantity (11–19%). Segment 1 and 2 households have more of their price sensitivity traceable to the quantity decision. For Segment 2, price sensitivity in quantity is even higher than price sensitivity in choice. In general, Segments 3 and 4 might be perceived as prone to brand switching and early buying but not stockpiling, whereas Segments 1 and 2 might be perceived as prone to stockpiling but not brand switching or early buying. These results corroborate those of Krishnamurthi and Raj (1991), who find that brand-loyal households have high purchase-quantity elasticities but low choice elasticities, whereas nonloyal households have low quantity elasticities but high choice elasticities.

CONCLUSION

Our purpose here is to develop a model of consumer decisions for brand choice, purchase incidence, and purchase quantity that simultaneously reveals latent response segments. We modeled the choice decision with multinomial logit, the incidence decision with nested logit, and the quantity decision by poisson regression. Consumer response segments across the three behaviors were obtained by simultaneously estimating the parameters of a probabilistic mixture model.

Application of our joint estimation approach to scanner panel data for the yogurt category reveals five distinct response segments. These segments differed in their overall price elasticity, ranging from a low of -0.74 to a high of -2.31 . Moreover, we found that the impact of price across the three behaviors also differed substantially by segment. In the two largest segments (79% of households, 63% of volume), price had a large impact on choice and incidence but a small impact on quantity. Price discounts induce these households to switch brands and buy early but have little effect on their stockpiling behavior. In two of the smaller segments (10% of households, 20% of volume), price had a smaller impact on choice and incidence but a greater impact on quantity. A fifth segment (11% of households, 17% of volume) had a low overall level of price sensitivity. An important implication of our results is that differences in response for one purchase behavior are not necessarily predictive of differences in response for other behaviors or for overall sales elasticities.

We also compared our joint estimation approach with a separate estimation approach. Although the separate estimation approach yields a better overall fit to the data, it specifies a much larger number of segments than the joint estimation approach (48 versus 5). Moreover, the estimated sizes for two-thirds of the segments (30 of 48) are less than or equal to 1% of the households. Therefore, use of the joint versus the separate estimation method involves a trade-off between the parsimony of the former and the ability to represent more combinations of response heterogeneity in the latter. The joint method also avoids the assumption that segmentation in response behavior is independent across choice, incidence, and quantity.

A practical problem for researchers attempting to use the separate estimation approach is likely to be the number of segments to retain. When using the joint approach, this problem can be handled with model selection criteria (i.e., BIC). When using the separate approach, the criterion is applied separately to choice, incidence, and quantity. Although each can be represented parsimoniously with several segments, the overall segmentation scheme easily becomes unwieldy, as the yogurt data show. The investigator must then either work with the total number of segments or subjectively prune them back to a more manageable number.

Our procedure also can be extended to determine brand-specific elasticities by segment. This procedure enables an analyst to identify brands that have the potential to increase category volume for the retailer (due to incidence and quantity) and increase brand share (due to choice). This should help managers identify opportunities for improved cooperation between manufacturers and retailers.

REFERENCES

- Ben-Akiva, Moshe and Steve R. Lerman (1985), *Discrete Choice Analysis*. Cambridge, MA: MIT Press.
- Bucklin, Randolph E. and Sunil Gupta (1992), "Brand Choice, Purchase Incidence, and Segmentation: An Integrated Approach," *Journal of Marketing Research*, 29 (May), 201–15.
- and James M. Lattin (1991), "A Two-State Model of Purchase Incidence and Brand Choice," *Marketing Science*, 10 (Winter), 24–39.
- Chiang, Jeongwen (1991), "A Simultaneous Approach to the Whether, What, and How Much to Buy Questions," *Marketing Science*, 10 (Fall), 297–315.
- Chintagunta, Pradeep (1993), "Investigating Purchase Incidence, Brand Choice and Purchase Quantity Decisions of Households," *Marketing Science*, 12 (Spring), 184–208.
- Gensch, Dennis H. (1985), "Empirically Testing a Disaggregate Choice Model for Segments," *Journal of Marketing Research*, 26 (May), 230–36.
- Guadagni, Peter M. and John D.C. Little (1983), "A Logit Model of Brand Choice Calibrated on Scanner Data," *Marketing Science*, 2 (Summer), 203–38.
- and ——— (in press), "When and What to Buy: A Nested Logit Model of Coffee Purchases," *Journal of Forecasting*, forthcoming.
- Gupta, Sunil (1988), "Impact of Sales Promotions on When, What, and How Much to Buy," *Journal of Marketing Research*, 25 (November), 342–55.
- Jain, Dipak C., Naufel J. Vilcassim, and Pradeep K. Chintagunta (1994), "A Random Coefficients Logit Brand-Choice Model Applied to Panel Data," *Journal of Business and Economic Statistics*, 12 (3), 317–28.
- Kamakura, Wagner A. and Gary J. Russell (1989), "A Probabilistic Choice Model for Market Segmentation and Elasticity Structure," *Journal of Marketing Research*, 26 (November), 379–90.
- Krishnamurthi, Lakshman and S.P. Raj (1988), "A Model of Brand Choice and Purchase Quantity Price Sensitivities," *Marketing Science*, 7 (Winter), 1–20.
- and ——— (1991), "An Empirical Analysis of the Relationship Between Brand Loyalty and Consumer Price Elasticity," *Marketing Science*, 10 (Spring), 172–83.
- Maddala, G. S. (1983), *Limited Dependent and Qualitative Variables in Econometrics*. Cambridge, UK: Cambridge University Press.
- McKelvey, Richard D. and William Zavoina (1975), "A Statistical Model for the Analysis of Ordinal Level Dependent Variables," *Journal of Mathematical Sociology*, 4 (1), 103–20.
- Neslin, Scott A., Caroline Henderson, and John Quelch (1985), "Consumer Promotions and the Acceleration of Product Purchases," *Marketing Science*, 4 (Spring), 147–65.