SUNIL GUPTA*

The effectiveness of a sales promotion can be examined by decomposing the sales "bump" during the promotion period into sales increase due to brand switching, purchase time acceleration, and stockpiling. The author proposes a method for such a decomposition whereby brand sales are considered the result of consumer decisions about *when, what,* and *how much* to buy. The impact of marketing variables on these three consumer decisions is captured by an Erlang-2 interpurchase time model, a multinomial logit model of brand choice, and a cumulative logit model of purchase quantity. The models are estimated with IRI scanner panel data for regular ground coffee. The results indicate that more than 84% of the sales increase due to promotion comes from brand switching (a very small part of which may be switching between different sizes of the same brand). Purchase acceleration in time accounts for less than 14% of the sales increase, whereas stockpiling due to promotion is a negligible phenomenon accounting for less than 2% of the sales increase.

Impact of Sales Promotions on When, What, and How Much to Buy

A question of continuing interest to marketing researchers and practitioners is how marketing mix variables affect consumers' purchase decisions and thus the sales of a brand. The interest is growing with the escalation in promotional expenditures. Though manufacturers can see a gratifying sales increase during a promotion period, a nagging question remains—is the increase in sales due to consumers switching from other brands or is the brand borrowing sales from the future as consumers advance their purchases in time or stockpile the product? This question can be answered by decomposing the sales "bump" during the promotion period into sales increases due to brand switching, due to purchase time acceleration, and due to stockpiling. This is accomplished by understanding the impact of sales promotions on consumer decisions of *when, what,* and *how much* to buy, which in turn determine the overall sales of a brand.

The mere fact that marketing variables affect consumers' purchase decisions is not new. Many studies have shown that price and sales promotions have a significant impact on consumers' brand choice, purchase time, and purchase quantity decisions (Blattberg, Eppen, and Lieberman 1981; Ehrenberg 1972; Guadagni and Little 1983; Kinberg, Rao, and Shakun 1974; Kuehn and Rohloff 1967; Massy and Frank 1965; Shoemaker 1979; Ward and Davis 1978; Wilson, Newman, and Hastak 1979). Several models have been developed to study the impact of marketing variables on brand choice (e.g., Blattberg and Jeuland 1981; Guadagni and Little 1983; Keon 1980; Lilien 1974), interpurchase time (e.g., Massy, Montgomery, and Morrison 1970; Raj, Staelin, and Mitchell 1977), purchase quantity (e.g., Blattberg et al. 1978; Blattberg, Eppen, and Lieberman 1981; Kunreuther 1973), brand choice and interpurchase time together (e.g., Jones and Zufryden 1980; Little and Guadagni 1986; Wagner and Taudes 1986), interpurchase time and purchase quantity together (e.g., Neslin, Henderson, and Quelch 1985), and brand choice and purchase quantity together (e.g., Krishnamurthi and Raj 1988).

Though brand sales are the result of consumers' decisions about when, what, and how much to buy, very little work has been done to capture these three consumer decisions at the same time for decomposing the effect of a marketing instrument into these three components. In this article the three consumer decisions are modeled to

*Sunil Gupta is Assistant Professor, John E. Anderson Graduate School of Management, University of California, Los Angeles.
The author thanks Donald R. Lehmann and four anonymous *JMR* reviewers for providing useful comments and Information Resources, Inc., for providing the data.

Journal of Marketing Research
Vol. XXV (November 1988), 342–55

decompose the sales "bump" during the promotion period into sales increase due to brand switching, purchase time acceleration, and stockpiling, which in turn helps in understanding the effectiveness of a sales promotion.[1]

The conceptual framework for the study is presented first, followed by the models. The data and the key explanatory variables then are described. Next the estimation procedure and results are reported. Finally, the contributions and limitations of this study are summarized.

CONCEPTUAL FRAMEWORK

We consider brand choice and purchase quantity decisions of a consumer to be conditional on the purchase time decision.[2] Further, brand choice and purchase quantity decisions are modeled and estimated separately. Though we make these simplifying assumptions for ease of estimation, we have at least some partial support for them.

First, in the stochastic modeling tradition, we model brand choice as conditional on purchase incidence. In other words, brand choice probability is modeled as a function of that week's market conditions in which the consumer decides to make a product purchase. Thus brand choice and interpurchase time can be modeled and estimated separately.

Second, Neslin, Henderson, and Quelch (1985) estimated the interpurchase time and purchase quantity jointly as well as separately. They found "the separate estimation produced parameter estimates not significantly different from the joint estimation results, while at the same time having smaller standard errors" (p. 155).

Third, separate modeling of brand choice and purchase quantity is similar in spirit to the aggregate modeling approach wherein brand sales are posited to be a product of category sales volume and brand market share (Naert and Leeflang 1978). Because in this study brand is defined as brand size and quantity is defined as the ounces of product bought, there are some obvious concerns about the separate estimation, especially if brand choice and purchase quantity are highly correlated. A simple test of the data, however, shows this association to be very weak.[3] Therefore, at least as a first-cut simplifying assumption, we can model and estimate brand choice and purchase quantity separately.

MODELS

The Brand Choice Model

The multinomial logit model has an extensive history of application in marketing (e.g., Carpenter and Lehmann 1985; Gensch and Recker 1979; Guadagni and Little 1983; Hauser 1978; Punj and Staelin 1978). It is appealing because it is based on a behavioral theory of utility, allows explanatory variables, and accounts for competition. We therefore use multinomial logit as our brand choice model.

$$(1) \qquad P_{ijn} = \frac{\exp(\mathbf{b}' \mathbf{X}_{ijn})}{\sum_{m=1}^{M} \exp(\mathbf{b}' \mathbf{X}_{imn})}, \qquad j = 1, \ldots, M \text{ brands,}$$

where:

P_{ijn} = probability that person i buys brand j on n^{th} purchase occasion and

$\mathbf{X}_{ijn}$ = explanatory variables for brand j and consumer i on the n^{th} purchase occasion, with response parameters **b**.

The assumptions and limitations of the logit model are discussed elsewhere (Guadagni and Little 1983; Lattin 1987; McFadden 1974).

The Interpurchase Time Model

Ehrenberg's (1959) negative binomial distribution (NBD) model has been a dominant model of interpurchase time. As Chatfield and Goodhardt (1973, p. 828) note:

> Although the NBD model has proved successful in a wide range of situations, . . . [its] theoretical assumptions have been questioned. . . . The assumption that individual consumers' purchases should be Poisson . . . can be criticized because of its implications concerning interpurchase times. If the assumption is true . . . then interpurchase times should follow an exponential distribution. The mode of the exponential distribution is at zero, but it is in fact improbable that a buyer is most likely to buy again *immediately*.

Several studies therefore have used an Erlang-2 distribution for interpurchase time because the mode of this

---

[1]Strictly, we are discussing decisions at the household level. "Consumer" and "household" are used interchangeably here.

[2]Therefore, brand choice and purchase quantity are modeled for every purchase occasion. Also, a brand in this study is defined as a "brand-size" (discussed subsequently).

[3]A simple cross-tabulation of brand sizes and purchase quantity categories shows they are statistically dependent. To test the strength of this association, we use lambda (Goodman and Kruskal 1954) and the uncertainty coefficient U (Brown 1975). A low value of these measures (lambda = .046, U = .053) suggests that the level of association between brand choice and purchase quantity is weak. Krishnamurthi and Raj (1988) indicate that the parameters may be biased and/or inefficient if the correlation between the errors of the two models is high. We estimate the squared error of the brand choice model as $\sum_{j=1}^{11} (a_j - p_j)^2$ where $a_j = 1$ if brand j is chosen, 0 otherwise, and p_j is the predicted probability of choosing brand j. (Note the errors are squared as $\Sigma(a_j - p_j) = 0$). Similarly, the squared error of the purchase quantity model is estimated as (actual − model-predicted quantity)^2. The correlation between these squared errors in the calibration dataset is −.02, suggesting that independent estimation of brand choice and purchase quantity will not introduce any significant biases or inefficiencies in estimating the parameters.

distribution is greater than zero (e.g., Chatfield and Goodhardt 1973; Herniter 1971; Morrison and Schmittlein 1981). These studies did not incorporate explanatory variables, though strong empirical evidence suggests that promotions can accelerate purchases in time (Shoemaker 1979; Ward and Davis 1978). We therefore propose that interpurchase time is Erlang-2 distributed where the mean of this distribution is a function of explanatory variables.[4] Thus the entire distribution of interpurchase time shifts left or right depending on such market conditions as price and promotion, as well as household inventory.

Hence the proposed interpurchase time model is

$$f_{iw}(t) = \alpha_{iw}^2 \, t \exp(-\alpha_{iw} \, t) \tag{2}$$

and

$$\alpha_{iw} = \exp(-\mathbf{c}'\mathbf{Y}_{iw}) \tag{3}$$

where:

$f_{iw}(t)$ = probability density of interpurchase time t for consumer i in week w,
α_{iw} = scale parameter of Erlang-2 distribution (an exponential function is used in equation 3 to ensure that α_{iw} is always positive), and
$\mathbf{Y}_{iw}$ = vector of explanatory variables that may affect the mean time, with response parameters **c**.

Because the mean of the Erlang-2 distribution is given by $2/\alpha_{iw}$, we get

$$\text{mean interpurchase time} = 2/\alpha_{iw} = 2\exp(\mathbf{c}'\mathbf{Y}_{iw}). \tag{4}$$

The Purchase Quantity Model

A simple model of purchase quantity (quantity defined as ounces of product bought by a household on a purchase occasion) would be a regression model where consumer's purchase quantity is represented as a linear function of exogenous variables (Neslin, Henderson, and Quelch 1985; Paull 1978). However, for many product categories, a regression model may not be appropriate because it assumes that the dependent variable is continuous. If quantity bought is discrete (e.g., 16, 32, or 48 oz of ground coffee), there may be strong biases if conventional regression is used. For example, regression may predict that a consumer will buy 20 oz of coffee. However, the consumer may actually buy 16 oz or 32 oz because coffee is available in discrete amounts only. This discrete nature of the dependent variable can be handled by the following procedure.

Let the response variable Q be an ordered categorical variable having $k = 1, \ldots, L$ possible values (here we are denoting 16 oz, 32 oz, etc. as category 1, 2, and so on). Further assume that the observed ordinal response variable Q has an underlying latent variable V on an interval scale. V can be thought of as the dependent variable of an ordinary regression model. In other words,

$$V = \boldsymbol{\beta}'\mathbf{Z} + \epsilon \tag{5}$$

where **Z** is a vector of explanatory variables (e.g., previous household inventory, interpurchase time) that are proposed to affect the consumer's purchase quantity decisions. The observed discrete purchase quantity (Q) for a consumer therefore can be represented as

$Q = 1$ | $Q = 2$ | . . . | $Q = k$ → V

θ_1 θ_2 θ_{k-1} θ_k

where:

$$Q = k \quad \text{if} \quad \theta_{k-1} < V < \theta_k \tag{6}$$
$$Q \le k \quad \text{if} \quad V \le \theta_k$$

and θ_k = unknown cutoff points with the ordinality constraint that

$$-\infty = \theta_0 < \theta_1 \le \theta_2 \le \ldots \le \theta_{L-1} < \theta_L = \infty.$$

Combining equations 5 and 6, it is clear that

$$P(Q \le k) = P(V \le \theta_k)$$
$$= P(\boldsymbol{\beta}'\mathbf{Z} + \epsilon \le \theta_k) = P(\epsilon \le \theta_k - \boldsymbol{\beta}'\mathbf{Z}).$$

Assuming the error ϵ has a logistic distribution, we get (replacing the subscripts for consumer i and purchase occasion n):

$$P(Q_{in} \le k) = \frac{\exp(\theta_k - \boldsymbol{\beta}'\mathbf{Z}_{in})}{1 + \exp(\theta_k - \boldsymbol{\beta}'\mathbf{Z}_{in})} \qquad k = 1, \ldots, L. \tag{7}$$

Thus,

$$P(Q_{in} = k) = P(Q_{in} \le k) - P(Q_{in} \le k - 1). \tag{8}$$

Equations 7 and 8 give the purchase quantity model, which is termed a cumulative logit or an ordered regression model (McCullagh 1980).

DATA

IRI (Information Resources, Inc.) scanner panel data for coffee were used for calibrating and validating the models. The dataset covers a panel of approximately 2000 households for a two-year period (1980–1982). It contains records of the complete purchase history of each household in the panel (e.g., household identification, brand bought, week and the minute when bought, quantity bought, store where bought, etc.). In addition, a stores file records weekly information on prices and promotions for all the coffee brands available in all the stores in the market.

To have a manageable data set, our analysis was restricted to the data for ground caffeinated coffee in one

[4]The Erlang-2 assumption was supported empirically by three different methods: method of moments, approximation to the maximum likelihood estimates (Dunn, Reader, and Wrigley 1983), and the Z-score method (Wheat and Morrison 1986). The average share parameter (r) for 100 households in the dataset was estimated as 1.78, 2.03, and 1.95 by the three methods, respectively. Further, for 92 of the 100 households, the null hypothesis of Erlang-2 could not be rejected by a Kolmogorov-Smirnoff test.

market—Pittsfield, MA. After elimination of non-continuously-reporting households and very light users (less than 10 purchases in two years), the data represented 395 households. Because of computer memory size limitations, a random sample of 100 households was selected from these 395 households. The purchase history of these 100 households then was combined with the store information so that the merged dataset had the information about competitive brands as well.

The data were available for IRI weeks 30 through 137. Weeks 30–59 were the precalibration period. This period's data were used to initialize certain variables (e.g., brand loyalty, household inventory). A reasonably long precalibration period ensured that these variables stabilized before they were used for calibration purposes. Weeks 60–100 were the calibration period and weeks 101–125 were the validation period. Weeks 126–137 were discarded and not used at all because no coffee purchase by a household in this period may be due to that household leaving the IRI panel. The calibration dataset contained 1526 purchase occasions (when coffee *was* bought) and 4211 purchase opportunities (when coffee *could* have been bought). The validation dataset contained 859 purchase occasions and 2535 purchase opportunities. Each week was assumed to provide a purchase opportunity for a consumer to buy coffee. (If a consumer made $n \geq 2$ coffee purchases in a week, that week was assumed to provide n purchase opportunities for that consumer.) Details related to data are provided in Appendix A.

Because different sizes of the same brand are not promoted at the same time, Guadagni and Little (1983, p. 213) note that ". . . different sizes of the same brand are clearly different products. . . . Therefore we model brand-sizes." We followed a similar approach whereby a brand was defined as brand-size. Thus Maxwell House 16 oz is a different brand from Maxwell House 32 oz. (Hereafter "brand" means brand-size, unless stated otherwise.) Further, different grinds (regular, drip, auto matic) of a brand were treated as one brand because they were priced and promoted together. Of the many brand-sizes available in the Pittsfield market, we selected 10 that account for about an 87.5% share of the market. These brands (and their market shares) are Martinson 16 oz (2.41% share), Hills Brothers 16 oz (8.57%), Folgers 16 oz (18.43%), Folgers 13 oz (2.84%), Chase & Sanborn 16 oz (2.59%), Maxwell House 16 oz (19.92%), Maxwell House 32 oz (3.36%), Master Blend 13 oz (6.12%), Chock-Full-O'Nuts 16 oz (16.58%), and private label brand 16 oz (5.63%). We also created an eleventh category of "all other brands" as the weighted sum of four major "other" brands.

VARIABLES

Variables for the Brand Choice Model

The following variables ($\mathbf{X}_{ijn}$, for consumer i, brand j, and purchase occasion n) are used for the brand choice model.

$X_{0,j}$ = brand-specific constant (0 or 1)
$X_{1,ijn}$ = regular brand price (cents/ounces)
$X_{2,ijn}$ = promotional price cut (cents/ounces)
$X_{3,ijn}$ = feature-or-display (between 0 and 1)
$X_{4,ijn}$ = feature-and-display (between 0 and 1)
$X_{5,ijn}$ = brand loyalty (between 0 and 1)
$X_{6,ijn}$ = size loyalty (between 0 and 1)

To capture any uniqueness of a brand that is not explained by any other variable, each brand has its own brand-specific constant.

IRI data contain weekly shelf price information for the brands. The shelf price consists of two components:

$$\text{shelf price} = \text{regular price} - \text{promotional price cut.}$$

Because no data are available on the regular price of brands, it is inferred from the shelf price. If the shelf price of a brand in a store remains constant for four or more consecutive weeks, it is assumed to be the regular price. Because different stores have different prices for a brand, regular price is weighted by an individual consumer's store share.[5] Variables X_2 through X_4 are weighted in the same way. Because different sizes of the product are available in the market, the regular price and promotional price cut are defined in cents/ounces (Guadagni and Little 1983).

Feature (in-store flier) and display (in store) are highly correlated (r = .704). To avoid this collinearity we construct two variables for each store s: X_{3s} = feature-or-display (1 if brand is either on feature or display in store s, 0 otherwise) and X_{4s} = feature-and-display (1 if brand is both on feature and display in store s, 0 otherwise). X_3 and X_4 then are constructed as the weighted averages of X_{3s} and X_{4s}, respectively, where the weights are the individual consumer's store shares as described before. The correlation between X_3 and X_4 is −.326.

Similar to Guadagni and Little (1983), we define brand and size loyalty as exponentially weighted averages of past purchases. Guadagni and Little estimated the smoothing constant as .875 for brand loyalty and .812 for size loyalty. However, Ortmeyer (1985) demonstrated their model's fit and parameter estimates are very robust to small changes (±.1) in the smoothing constant. Hence, instead of estimating, we prespecify the constants as .8. A sensitivity analysis with smoothing constants as .7 and .9 corroborated Ortmeyer's findings. (Recently, in a different modeling context, Lattin 1987 prespecified a similar smoothing constant as .7 and found results to be insensitive to small changes in this constant.) Brand and size loyalty are initialized as suggested by Guadagni and Little (1983). For example, for the first purchase occasion of consumer i, $X_{5,ij}$ is set equal to the smoothing constant value of .8 if brand name of brand-

[5]Ideally, we would compute store shares for a household as the percentage of *shopping trips* it makes to a particular store. However, because of lack of such information, the percentage of *coffee purchases* made in a store is used.

size j was bought, otherwise .2/(number of brand names − 1). This step ensures that the sum of loyalties across brand names always equals 1 for a consumer. Thus two brand-sizes with the same brand name (e.g., Folgers 13 oz and Folgers 16 oz) have the same value for brand loyalty.

For all variables except the brand-specific constants, the response coefficients are assumed to be the same for all brand sizes. Hence the model is relatively parsimonious.

Variables for the Interpurchase Time Model

The following variables ($\mathbf{Y}_{iw}$, for consumer i, week w) are used for the interpurchase time model.

$Y_{1,i}$ = average interpurchase time of a household (weeks)
$Y_{2,i(w-1)}$ = an estimate of household's previous week's product inventory (ounces)
$Y_{3,iw}$ = promotional price cut for the product category (cents/ounces)
$Y_{4,iw}$ = feature-or-display for the product category (between 0 and 1)
$Y_{5,iw}$ = feature-and-display for the product category (between 0 and 1)
$Y_{6,iw}$ = product price (cents/ounces)

The average interpurchase time of a household is based on its purchase frequency during the calibration period. The IRI dataset indicates the week and the minute of the week when a purchase is made. Hence the precision of interpurchase time is up to a minute.

Following Neslin, Henderson, and Quelch (1985), we estimate weekly product inventory as

$$\begin{matrix}\text{estimated} \\ \text{inventory for} \\ \text{week } w-1\end{matrix} = \begin{matrix}\text{estimated} \\ \text{inventory for} \\ \text{week } w-2\end{matrix} - \begin{matrix}\text{estimated} \\ \text{weekly} \\ \text{consumption}\end{matrix} + \begin{matrix}\text{quantity bought} \\ \text{in week } w-1, \\ \text{if any.}\end{matrix}$$

Further, in line with Neslin, Henderson, and Quelch (1985), for each household the initial inventory is estimated as its average purchase quantity and the consumption rate as its average buying rate over the calibration period. Because consumption data are not available, the weekly consumption rate is assumed to be constant over time. If a consumer does not buy coffee for a long time, the inventory value may be negative. Because inventory cannot be negative, we set it to zero in such cases.

The decision of when to buy can be influenced by the price and promotion of not only one brand, but of all the brands in the market. In other words, product category promotion rather than just one brand's promotion is likely to influence the interpurchase time. We create an index for product promotion as a weighted sum of individual brands' promotions. Market shares typically are used as the weights to create such an index (Neslin and Shoemaker 1983). We use individual-level weights, taking each household's share of brand purchases in the calibration period as the weight for that brand and that household. Variables Y_3 through Y_6 are created by using this weighting scheme. A similar weighting scheme was used by Krishnamurthi and Raj (1985).

Variables for the Purchase Quantity Model

For the purchase quantity model, quantity is defined as ounces of regular ground coffee a consumer buys on any one purchase occasion. Eight discrete categories are selected, based on the frequency plot of quantity bought by all consumers. The explanatory variables ($\mathbf{Z}_{in}$, for consumer i, purchase occasion n) follow.

$Z_{1,i}$ = average purchase quantity of a household (ounces)
$Z_{2,i(n-1)}$ = an estimate of household's inventory on previous purchase occasion (ounces)
$Z_{3,in}$ = time between previous and current purchase (weeks)
$Z_{4,in}$ = promotional price cut for the product category (cents/ounces)
$Z_{5,in}$ = feature-or-display for the product category (between 0 and 1)
$Z_{6,in}$ = feature-and-display for the product category (between 0 and 1)
$Z_{7,in}$ = product price (cents/ounces)
$Z_{8,i}$ = household or family size

Similar to average interpurchase time, average purchase quantity of a household is derived from its purchase history during the calibration period. Household product inventory on previous purchase occasion is estimated as

$$\begin{matrix}\text{estimated} \\ \text{inventory on} \\ n-1 \text{ purchase}\end{matrix} = \begin{matrix}\text{estimated} \\ \text{inventory on} \\ n-2 \text{ purchase}\end{matrix} - \begin{pmatrix}\text{estimated} \\ \text{weekly} \\ \text{consumption}\end{pmatrix} * \begin{pmatrix}\text{time between} \\ n-2 \text{ and } n-1 \\ \text{purchase}\end{pmatrix} + \begin{matrix}\text{quantity} \\ \text{bought on} \\ n-1 \text{ purchase.}\end{matrix}$$

Note that we use weekly inventory for the interpurchase time model because it focuses on each week as a purchase opportunity. However, purchase quantity focuses on a purchase occasion as an observation.

The promotion and price indices used are similar to those used for the interpurchase time model. Notice that we use consumer's brand preferences as weights to construct the promotion indices. Hence, when a consumer's favorite brand is on promotion, the promotion indices have a higher value. Because these promotion indices are expected to have a positive impact on the consumer's purchase quantity, we expect a consumer to stockpile if his or her favorite brand is on promotion. Finally, a demographic variable, family size, is included that is expected to be correlated positively with consumption (Krishnamurthi and Raj 1985) and hence with quantity.

Table 1 is a summary of the hypothesized signs of the parameters.

ANALYSIS AND RESULTS

Each model was estimated separately by the maximum likelihood procedure. Eighteen parameters were esti-

Table 1
HYPOTHESIZED SIGNS OF PARAMETERS

Variable	*Expected sign of parameter*	*Comment*
Brand choice model		
Regular price ($X_{1,ijn}$)	−	Higher the price, lower the choice probability
Price cut ($X_{2,ijn}$)	+	Promotion enhances brand's value, hence its probability of choice
Feature-or-display ($X_{3,ijn}$)	+	
Feature-and-display ($X_{4,ijn}$)	+	
Brand loyalty ($X_{5,ijn}$)	+	Higher the loyalty, higher the choice probability
Size loyalty ($X_{6,ijn}$)	+	
Interpurchase time model		
Average purchase time ($Y_{1,i}$)	+	
Inventory ($Y_{2,i(w-1)}$)	+	Higher inventory makes consumers wait longer
Price cut ($Y_{3,iw}$)	−	Good deal may make a consumer buy early
Feature-or-display ($Y_{4,iw}$)	−	
Feature-and-display ($Y_{5,iw}$)	−	
Regular price ($Y_{6,iw}$)	+	Product purchase may be delayed because of high price
Purchase quantity model		
Average purchase quantity (Z_{1-i})	+	
Inventory ($Z_{2,i(n-1)}$)	−	Higher inventory → buy less
Interpurchase time ($Z_{3,in}$)	+	Longer wait between purchases → buy more
Price cut ($Z_{4,in}$)	+	Good deal may make a consumer stockpile
Feature-or-display ($Z_{5,in}$)	+	
Feature-and-display ($Z_{6,in}$)	+	
Regular price ($Z_{7,in}$)	−	Negative price elasticity
Family size ($Z_{8,i}$)	+	Consumption increases with family size

mated for the brand choice model (10 parameters for brand-specific constants plus eight response parameters), seven parameters for the interpurchase time model (one constant plus six response parameters), and 15 parameters for the purchase quantity model (seven cutoff points θ_k plus eight response parameters). The LOGIST program of SAS was used to estimate the purchase quantity model and the PAR program of BMDP was used to estimate the brand choice and interpurchase time models. Various nested versions of the models were estimated to see the improvement in the model by adding variables (the decision rule for adding variables was to choose that variable contributing maximally to model fit). Model improvement and quality of fit were evaluated by using a likelihood ratio index ρ^2 (Hauser 1978), adjusted likelihood ratio index $\bar{\rho}^2$ (Ben-Akiva and Lerman 1985), and chi square (Ben-Akiva and Lerman 1985).

Discussion of Nested Models and Parameter Estimates

Brand choice model. The market share model, a model with only brand-specific constants, is taken as the null model (MNL_0), $\rho^2 = 0$. The results (Table 2) show that all significant parameters have the expected sign. Also, the explanatory power of the proposed model is very high, as is evident from an excellent fit of the full model MNL_5 ($\rho^2 = .554$). In other words, market variables such as price and promotion and consumer characteristics such as brand and size loyalty help substantially in explaining consumers' brand choice decisions.[6] Further, our results are consistent with Guadagni and Little's (1983, p. 221) finding that "brand and size loyalty [are the] most important [variables]."[7] High t-values and substantial improvement in model fit suggest that promotional variables also have a strong role in consumers' brand choice decisions. The addition of price cut to MNL_3 changes the coefficients of feature-and-display (F-and-D) and feature-or-display (F-or-D) substantially, suggesting a possible interaction between these variables. These in-

[6]One could argue that brand and size loyalty cannot be taken as explanatory variables because they capture the heterogeneity of consumers and their past purchase history but do not explain either of these factors. Consequently we may want to include brand and size loyalty variables in the null model. MNL_1 therefore becomes the null model instead of MNL_0. Comparing the nested models with MNL_1, we find that the quality of fit of the models is still very high (ρ^2 for MNL_5 is .421). This finding further strengthens the notion that such marketing variables as price and promotion have a significant role in consumers' brand choice decisions.

[7]As we are using cross-sectional and time series data, the loyalty variables capture both the heterogeneity and time-varying effects. A possible way to separate these two effects is to split the loyalty variables into two parts, base loyalty and dynamic loyalty. Base loyalty is a constant for a given household and reflects the average share of brand purchases during the calibration period. It therefore captures the heterogeneity effect. Dynamic loyalty is the difference between loyalty (as currently defined) and base loyalty. It therefore captures the time-varying effect. Results for the full model (MNL_5) show that almost all the brand and size loyalty effect is due to consumer heterogeneity (a model with base loyalties improves only marginally, by ρ^2 of .009, when dynamic loyalty variables are added). The parameters (and t-values in parentheses) follow.

	Loyalty	*Base loyalty*	*Dynamic loyalty*
Brand loyalty	4.6	5.8	−1.0
	(28.3)	(24.0)	(−2.3)
Size loyalty	2.2	3.5	−1.4
	(9.8)	(10.0)	(−2.0)

Interestingly, dynamic loyalties have negative parameters, which seems to suggest that a consumer who buys more than his or her usual share of a brand (or size) is expected to switch away from it (bringing the brand share down to its average). In spirit, this effect is similar to the "balanced choice behavior" discussed by Lattin (1987). Clearly these are preliminary findings and need more research.

Table 2
BRAND CHOICE MODEL: MULTINOMIAL LOGIT

	Parameter estimates (t-values)					
	MNL₀[a]	*MNL₁*	*MNL₂*	*MNL₃*	*MNL₄*	*MNL₅*
Brand loyalty		3.848	4.373	4.340	4.785	4.555
		(30.14)	(28.77)	(26.91)	(29.91)	(28.26)
Size loyalty		2.207	2.970	2.098	2.096	2.209
		(10.60)	(12.79)	(9.59)	(9.38)	(9.80)
Feature-and-display			4.068	2.764	3.868	3.543
			(36.89)	(24.16)	(26.27)	(22.81)
Feature-or-display			2.143	1.528	1.927	1.658
			(16.96)	(11.06)	(14.72)	(10.63)
Price cut				.716	1.157	1.147
(cents/oz)				(16.14)	(15.61)	(14.85)
F-and-D * price cut					−.802	−.757
					(−8.41)	(−7.68)
F-or-D * price cut					−.398	−.360
					(−3.48)	(−.317)
Regular price						−.117
(cents/oz)						(−1.72)
−LL[b] (n = 1462)	3020.33	2325.40	1532.57	1378.61	1354.29	1346.03
ρ^2[c]	—	.230	.493	.544	.552	.554
	[—]	[—]	[.341]	[.407]	[.418]	[.421]
$\bar{\rho}^2$[c]	—	.229	.491	.542	.549	.552
	[—]	[—]	[.340]	[.406]	[.415]	[.419]
χ^2[d]	—	1389.86	1585.66	307.92	48.64	16.52
d.f	—	2	2	1	2	1

[a]Brand-specific constants not reported.
[b]−LL = minus log likelihood.
[c]ρ^2 and $\bar{\rho}^2$ of nested models are with respect to null model MNL_0. Numbers in brackets are ρ^2 and $\bar{\rho}^2$ if MNL_1 is the null model.
[d]χ^2 for each model is with respect to the previous nested model; e.g., χ^2 for model MNL_4 is with respect to MNL_3.

teraction terms (price cut * F-and-D and price cut * F-or-D) are therefore added in MNL_4. The negative signs of the parameters suggest a possible overlap or substitutability among different promotional instruments. Notice that the inclusion of interaction terms stabilizes the F-and-D and F-or-D parameters (parameter values are now close to those in MNL_2 and t-values are higher than those in MNL_3). Finally, regular price does not have a strong role in consumers' brand choice decisions and is only marginally significant.

Interpurchase time model. ERL_0, with no household or marketing variables, is taken as the null model. The results (Table 3) show that all significant parameters have the expected sign. As expected, average interpurchase time of a household is the most important variable in explaining when that household buys coffee. This household-specific variable, which remains constant over time for a household, accounts for population heterogeneity in purchase timing. F-and-D and F-or-D are the next most important variables, suggesting that feature and display are likely to accelerate consumers' purchases in time. The parameter estimate of household inventory is also significant, confirming Neslin, Henderson, and Quelch's (1985) finding that a household is likely to wait longer if its previous product inventory is large. Finally, counter to our hypothesis, addition of price cut (ERL_4) and regular price (ERL_5) to the model is not useful because both the parameters and the improvement in the model are not significant. Hence, ERL_3 is the most parsimonious model and is used in further analysis.

These results indicate that a large portion of the variance in interpurchase time remains unexplained (ρ^2 for ERL_3 is .067 only). Further, most of this variation is accounted for by households' average interpurchase time. Thus marketing variables such as price and promotion do not have a large influence on consumers' purchase time decisions. Though the impact of feature and/or display is statistically significant, price cut and regular price have no significant impact. This finding suggests that a consumer who is not planning to buy coffee in a given week may not check the prices or price discounts on coffee brands unless his or her attention is attracted to them through feature and/or display.

Purchase quantity model. CL_0, with no household or marketing variables, is taken as the null model. The results (Table 4) show that all significant parameters have the expected sign. Household's average purchase quantity is the most important variable in explaining the quantity of coffee a household buys on any purchase occasion. As in the case of the interpurchase time model, this variable accounts for population heterogeneity in purchase quantity. Marketing variables (F-and-D, price cut, regular price, F-or-D) and household-specific variables (family size, inventory, interpurchase time) are added to the model stepwise as shown in Table 4. From the t-values of the parameters and the improvement of

Table 3
INTERPURCHASE TIME MODEL: ERLANG-2

	Parameter estimates (t-values)					
	ERL_0	*ERL_1*	*ERL_2*	*ERL_3*	*ERL_4*	*ERL_5*
Constant	.109	−.474	−.520	−.311	−.311	−.300
	(6.53)	(−11.05)	(−11.20)	(−6.12)	(−6.06)	(−1.03)
Average interpurchase		.219	.222	.210	.211	.208
time (week)		(14.60)	(14.73)	(13.54)	(13.53)	(13.28)
Feature-and-display			−1.274	−1.285	−1.217	−1.194
			(−11.00)	(−11.42)	(−9.06)	(−8.81)
Feature-or-display			−.588	−.612	−.566	−.546
			(−4.69)	(−4.82)	(−4.39)	(−4.10)
Inventory				.001	.001	.001
(oz)				(2.74)	(2.77)	(2.75)
Price cut					−.037	−.050
(cents/oz)					(−.83)	(−1.11)
Regular price						−.0003
(cents/oz)						(−.02)
−LL (n = 4160)	2556.6	2454.0	2387.2	2384.1	2383.5	2383.5
ρ^2	—	.040	.066	.067	.068	.068
$\bar{\rho}^2$	—	.040	.065	.066	.066	.065
χ^2	—	205.2	133.62	6.2	1.2	0
d.f.	—	1	2	1	1	1

the nested models, it is clear that CL_5 is the most suitable model (ρ^2 = .115).

Comparing CL_5 with CL_1, we notice that most of the variation in the purchase quantity is accounted for by consumers' average purchase quantity. Thus marketing variables have a very small influence on how much coffee a consumer buys. Apart from average purchase quantity of a household, variables that do have some impact on consumers' purchase quantity decisions are regular price of coffee, promotional variables such as price cut and F-and-D, and household variables such as family size.

The threshold constants (θ_k) for model CL_5 are given at the bottom of Table 4. The interpretation of these constants follows.

The threshold constants (θ_k) for model CL_5 are given at the bottom of Table 4. The interpretation of these constants follows.

Q = 13 oz 16 26 32 48 64 96 176 oz → V

θ_1 θ_2 θ_3 θ_4 θ_5 θ_6 θ_7
−2.46 1.19 1.30 3.47 4.42 5.20 6.36

Note: some quantities (e.g. 39 oz, 80 oz, 112 oz) have very few observations and hence are not assigned a separate category. They are instead combined with another category (e.g. 39 oz is combined with 48 oz category).

As indicated before, though observed quantity Q is discrete, we assume that a latent variable V is present on an interval scale so that

$$V = \boldsymbol{\beta}'\mathbf{Z} + \epsilon, \text{ where } \mathbf{Z} = \text{explanatory variables, and}$$

$$Q = \text{category } k \text{ if } \theta_{k-1} < V < \theta_k$$

This implies that if $\boldsymbol{\beta}'\mathbf{Z} < -2.46$ then Q = 13 oz, if $-2.46 < \boldsymbol{\beta}'\mathbf{Z} < 1.19$ then Q = 16 oz, and so on.

Notice that θ_2 and θ_3 are close. Because very few purchases of 26 oz are made by the consumers in this dataset, this finding suggests that the two categories, 16 oz and 26 oz, can be combined into one category.

Model Validation

Significant parameters with the expected signs and good fit (especially for the brand choice model) provide some validity for the proposed models. A further test of their accuracy is to examine their ability to predict correctly brand sales volume in the future period. A 25-week period (IRI weeks 101–125) of the same 100 households is used for validation.

Actual prices and promotions of all brand-sizes are used (i.e., we are not attempting to predict what marketing decisions will be made) in conjunction with the parameter estimates obtained from model calibration to predict whether a product purchase will be made by a consumer in a given week. If a product purchase is predicted, the market conditions of that week are used to predict which brand and how much will be bought. The brand sales then are aggregated across consumers to get weekly sales volume.

Note that some of the independent variables (e.g., brand loyalty) need updating based on the previous purchase history of a consumer. Like Guadagni and Little (1983), we use a Monté Carlo procedure whereby a brand purchase is simulated on the basis of the brand choice probabilities as predicted by the model. This synthetic purchase history of a consumer is used to update his or her brand loyalty and size loyalty variables. To avoid any distortion from a single unlikely choice generated by the Monté Carlo procedure, we repeat the procedure six times

Table 4
PURCHASE QUANTITY MODEL: CUMULATIVE LOGIT

	Parameter estimates (t-values)								
	CL_0	CL_1	CL_2	CL_3	CL_4	CL_5	CL_6	CL_7	CL_8
Average purchase		.107	.107	.108	.112				
quantity (oz)		(16.04)	(15.89)	(15.98)	(15.69)				
			.106	.112					
			(15.84)	(15.68)					
			.107	.112					
			(15.86)	(15.69)					
Feature-and-			1.454	.966	.890	.944	.958	1.014	1.104
display			(4.54)	(2.67)	(2.44)	(2.58)	(2.61)	(2.63)	(2.63)
Price cut				.321	.336	.316	.310	.292	.292
(cents/oz)				(2.86)	(2.98)	(2.79)	(2.73)	(2.43)	(2.43)
Regular price					−.164	.172	−.180	−.175	−.175
(cents/oz)					(−3.09)	(−3.22)	(−3.36)	(−3.23)	(−3.21)
Family size						.077	.073	.075	.075
						(1.99)	(1.89)	(1.92)	(1.91)
Inventory (oz)							−.002	−.002	−.002
							(−1.71)	(−1.67)	(−1.67)
Feature-or-								.185	.185
display								(.47)	(.47)
Interpurchase									.0003
time (weeks)									(.01)
−LL (n = 1526)	1787.9	1608.6	1592.8	1588.8	1583.9	1581.9	1580.5	1580.4	1580.4
ρ^2	—	.100	.109	.111	.114	.115	.116	.116	.116
$\bar{\rho}^2$	—	.100	.108	.110	.112	.112	.113	.112	.112
χ^2	—	358.6	31.6	8.0	9.8	4.0	2.8	.2	0
d.f.	—	1	1	1	1	1	1	1	1
Threshold constants for CL_5	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6	θ_7		
Parameter estimate	−2.463	1.193	1.304	3.473	4.424	5.204	6.361		
(t-value)	(−2.92)	(1.43)	(1.56)	(4.10)	(5.13)	(5.91)	(6.87)		

and use the average of these six iterations (Lattin 1987; Ortmeyer 1985).

The models track weekly sales volume for each of the 11 brands fairly accurately (for most brands, the correlation between the actual and the predicted weekly brand sales is .8 or higher). The predictions are less accurate for the low market share brands, for the private label brand, and for "all other brands." Large market share brands are tracked very accurately because estimates are likely to be based predominantly on the data from these large brands. The predictive accuracy of the model is especially impressive because, unlike Guadagni and Little (1983), we are predicting not only brand share but also purchase timing and purchase quantity.

To evaluate further the predictive quality of the models, it is instructive to compare them with some naïve models. Three naïve disaggregate models are used, each consisting of a brand choice component, a purchase time component, and a purchase quantity component.

Naïve model 1. Here each consumer makes a purchase such that his or her interpurchase time equals his or her *average* interpurchase time and his or her purchase quantity equals his or her average purchase quantity in the calibration period. Consumer's brand share in the calibration period is used to simulate his or her brand choice on a given purchase occasion. Notice that this model accounts for consumer heterogeneity in interpurchase time, purchase quantity, and brand preferences.

Naïve model 2. For each consumer, the interpurchase time and purchase quantity are modeled as in naïve model 1. The brand choice also is modeled similarly except that the vector of brand shares is updated (like brand loyalty, with smoothing constant = .8) after each purchase simulation. Thus, in addition to consumer heterogeneity, this model accounts for changes over time in consumers' brand preferences.

Naïve model 3. The interpurchase time and purchase quantity for each consumer are modeled as in naïve models 1 and 2. The logit model (with all the explanatory variables) is used to predict brand choice. Comparison of this naïve model with our proposed models therefore shows the improvement in brand sales predictions due to the addition of the proposed interpurchase time and purchase quantity models.

These three naïve models are used to predict when, what, and how much a consumer would buy in the validation period. The brand sales then are aggregated across consumers to get the weekly brand sales predictions.

The predictive quality of these naïve models is compared with that of our proposed models by Theil's (1967) inequality coefficient U, as modified by Naert and Weverbergh (1981) for multiple brands. U ranges from 0 to 1, where smaller values mean better predictions. Predictive quality of each component of the model (i.e., choice, timing, and quantity), as well as the three components taken together, follows.

	U statistics for			
	Naïve model 1	*Naïve model 2*	*Naïve model 3*	*Proposed model*
Brand choice	.418	.410	.341	.281
Interpurchase time	.145	.145	.145	.112
Purchase quantity	.173	.173	.173	.124
All three together	.480	.469	.383	.308

As expected, the largest improvement in the overall predictive quality of our model is due to the brand choice component. Interestingly, though the brand choice model is the same in naïve model 3 and the proposed model, they differ in predictive quality. The reason is that brand choice is conditional on purchase time. Therefore a bad model of interpurchase time not only reduces its own predictive quality but also adversely affects that of the brand choice component. The results show that the proposed set of models are superior to the naïve models in predictions.

Decomposing the Effect of Promotion

One of the key questions this study set out to address is whether a promotional sales increase is due to consumers switching brands, accelerating purchases, or stockpiling the product. Estimation results of the brand choice, interpurchase time, and purchase quantity models suggest that marketing variables, specifically promotion variables, have their greatest influence on consumers' brand choice behavior. These variables have higher relative contribution to ρ^2 for the brand choice model than for the interpurchase time or purchase quantity model.

Though this "eyeballing" method gives some useful indications, it can be considered only a first step. Because of the different model structures used, the parameter estimates are not directly comparable across the three models. An elasticity analysis is performed to assess the relative impact of sales promotion on consumers' brand choice, interpurchase time, and purchase quantity decisions.

Method

Recall two assumptions of our modeling framework. First, brand choice and purchase quantity are conditional on purchase time. Second, brand choice and purchase quantity decisions are independent. These assumptions enable us to write the sales of brand j in any time interval 0 through t as a separable function.

$$Q_j = p_j Q p_t \tag{9}$$

where:

Q_j = sales of brand j in time interval $0\text{–}t$,

p_j = brand j's choice probability (conditional on product purchase),

Q = product category sales, and

p_t = probability of product purchase in time $0\text{–}t$ (i.e., purchase incidence probability).

If D_j is the level of deal or sales promotion of brand j during the time period $0 - t$, then by using equation 9, the deal elasticity of brand j is given by the chain rule for the product of functions (cf. Cooper and Nakanishi 1988) as

$$\eta_j = \frac{D_j}{Q_j}\frac{\delta Q_j}{\delta D_j} = \frac{D_j}{p_j}\frac{\delta p_j}{\delta D_j} + \frac{D_j}{p_t}\frac{\delta p_t}{\delta D_j} + \frac{D_j}{Q}\frac{\delta Q}{\delta D_j} \tag{10}$$

or

$$\eta_j = \eta_{BC(j)} + \eta_{PT} + \eta_{PQ} \tag{11}$$

where:

η_j = total elasticity of brand j (with respect to deal),

$\eta_{BC(j)}$ = brand choice elasticity for brand j,

η_{PT} = purchase time elasticity with respect to deal of brand j, and

η_{PQ} = purchase quantity elasticity with respect to deal of brand j.

Using the proposed models, we obtain $\eta_{BC(j)}$, η_{PT}, and η_{PQ} as a function of model parameters, brand j deal, and other explanatory variables. Mathematical derivations and expressions for the elasticities are given in Appendix B. Using these expressions and the mean values for all variables during the calibration period, we obtain the brand choice, purchase time, and purchase quantity elasticities.[8] The elasticity analysis is done for three promotional variables: feature-and-display (F-and-D), feature-or-display (F-or-D), and promotional price cut.

A simple example clarifies these elasticity calculations. For example, Appendix B indicates that

$$\eta_{PT} = \alpha^2 c s_j D_j / [1 + \alpha - \exp(\alpha)]$$

where:

$\alpha = \exp(-\mathbf{c}'\mathbf{Y})$, $\mathbf{Y}$ = explanatory variables,

D_j = deal of brand j with response parameter c, and

s_j = market share of brand j.

Using the parameters ($\mathbf{c}$) of the interpurchase time model (Table 3) and the mean values of the explanatory variables ($\mathbf{Y}$) during the calibration period, we get α. Now, to obtain η_{PT} for, say, Folgers 16 oz with respect to its F-and-D, we use its mean F-and-D value (D_j = .086), its market share during the calibration period (s_j = .232),

[8]Because the relationship of deal with choice, time, and quantity is nonlinear, the elasticities will be different at different values of these variables. However, for ease of exposition, we focus on short-term, point elasticities at the mean values of the variables.

and the F-and-D parameter from Table 3 ($c = -1.285$). Substituting these values in the preceding equation, we get η_{PT} for Folgers 16 oz with respect to its F-and-D as .0338. Similarly, we get η_{BC} and η_{PQ} for this brand as .2101 and .0042 respectively. Using equation 11, total elasticity is

$$\eta_j = \underset{(84.7\%)}{.2101} + \underset{(13.6\%)}{.0338} + \underset{(1.7\%)}{.0042} = \underset{(100\%)}{.2481}$$

Results

The results of elasticity analysis for all the brands and for all three promotional instruments (F-and-D, F-or-D, and promotional price cut) show that *of the total sales increase due to promotion, more than 84% is accounted for by brand switching, 14% or less by purchase time acceleration, and less than 2% by stockpiling.*

This finding indicates that in the regular ground coffee market, promotions are very effective in drawing consumers from competitive brands. Because brand is defined as brand-size, part of this switching may be between different sizes of the same brand. However, a simple switching matrix for the calibration data shows that size switching due to promotion accounts for less than 4% of the total brand name and size switching. This finding is not surprising because in our data there are few brands with the same name but different sizes—Folgers 13 oz and 16 oz and Maxwell House 16 oz and 32 oz. Further, Folgers 13 oz and Maxwell House 32 oz account for about 6% of the market. Thus most of the sales increase due to promotion comes from switching from competitive brands.

Almost all (more than 98%) of the sales increase due to price cut comes from brand switching. The reason is that, unlike F-and-D and F-or-D, price cut does not affect consumers' purchase time decisions (see Table 3). Stockpiling remains a negligible phenomenon.

Promotional sales increase due to purchase time acceleration and stockpiling cannot be considered as true incremental sales because at least some of these consumers would have bought the promoted brand in the future anyway. For all the brands, the sales increase due to purchase acceleration is less than 14% and that due to stockpiling is less than 2%.

The small effect of promotion on stockpiling can be ascribed to three possible reasons—consumer perception that stockpiling unusual amounts of coffee may destroy its freshness (a key attribute of coffee quality), storage constraints due to the large volume of a coffee can, and high promotion intensity in the marketplace ensuring that some brand is almost always on promotion. Hence stockpiling could be a much more important factor in other product categories such as tunafish (for which storage is not a problem) or paper towels (for which freshness is not important). This possibility should be investigated in future research.

CONCLUSIONS, IMPLICATIONS, AND LIMITATIONS

Our study examines the impact of promotions on consumer decisions of when, what, and how much to buy. The analysis for coffee data indicates that more than 84% of the sales increase due to promotion comes from brand switching (a very small part of which may be switching between different sizes of the same brand). Purchase acceleration in time accounts for less than 14% of the sales increase, whereas stockpiling accounts for less than 2% of the sales increase due to promotion. These results can be very useful to managers in understanding the effectiveness of a sales promotion. The proposed method also can be used to compare the effectiveness of alternative promotional offerings under various competitive conditions so as to determine the most suitable and effective promotion.

Some of the limitations of the study are also potential ideas for future research. First, the fact that when, what, and how much to buy are interrelated decisions for a consumer suggests that the "ideal" way to model these three decisions is to construct a large simultaneous system. However, because of the modeling and estimation complexity of such a simultaneous system, we made some simplifying assumptions. Future research therefore should attempt to relax these assumptions.

Second, we ignore the lead effect of promotions. Possibly consumers' purchase decisions are affected by their anticipation of future promotions (Doyle and Saunders 1985).

Third, the study addresses the issue of short-term impact of promotions, for which they are found to be beneficial. However, several studies suggest that past promotional purchases are likely to have a negative impact on repurchase probability of the brand. In other words, a promotion may have a negative long-term effect. More research is needed in this area.

Fourth, a key variable not considered in the study is the use of coupons. Because coupons can have a great impact on consumers' purchase decisions, their omission can lead to specification errors. However, inclusion of coupons in the model requires data about coupon availability (i.e., when a coupon was made available to a consumer for use) whereas most datasets contain information about coupon redemption.

Fifth, the brand choice model assumes that even a 5-cents-off promotion will change the brand choice probabilities of a consumer. However, it is more intuitive to assume that unless the promotional offering is beyond a threshold level, it will have no effect on choice probabilities. Introducing the concept of threshold would be a useful next step.

In spite of its limitations, the study provides a useful method for decomposing the sales "bump" during promotion into sales increase due to brand switching, purchase time acceleration, and stockpiling. Much work remains to be done, but the study takes a significant step toward better understanding of the complex and interrelated decisions made by consumers in a changing market environment.

APPENDIX A
DATA-RELATED ISSUES

A few households purchased multiple brands on a single purchase occasion. As there were very few occurrences of this problem (less than 1.5% of all purchase occasions), it was resolved by randomly assigning any one of the multiple brands to the purchase record.

Light users (less than 10 purchases in two years) were eliminated from the data because variables needing initialization (e.g., brand loyalty) or using household averages (e.g., average purchase quantity) would not be stable for those households. These eliminated households accounted for 1.7% of the purchase occasions and 1.5% of total coffee sales in ounces.

Of the 395 usable households in the dataset, only 100 (randomly selected) households were used for analysis because of computer memory size limitations. However, the sample size is large enough to provide robust results.

The calibration dataset for the brand choice and purchase quantity models contained 1526 purchase *occasions* (i.e., when coffee was bought). Because of missing values of certain observations, 1462 observations were finally used for calibrating the brand choice model. The validation dataset contained 859 purchase occasions.

The unit of observation for the interpurchase time model is purchase *opportunity* (i.e., when coffee could have been bought) instead of purchase *occasion*. In the calibration period (41 weeks), the number of purchase opportunities for the 100 households should be 41 × 10 = 4100. However, some households made more than one purchase in certain weeks. Hence the total number of purchase opportunities in the calibration period was 4211. Because of missing values for certain observations, 4160 observations were finally used for calibrating the interpurchase time model. Similarly, the number of purchase opportunities in the validation period was 2535.

APPENDIX B
ELASTICITY ANALYSIS

We use the following expressions for brand choice elasticity, purchase time elasticity, and purchase quantity elasticity. The total promotional elasticity of a brand can be obtained as a sum of these three elasticities (see equation 11).

Elasticities are calculated at the mean value of the variables. We therefore drop subscripts i (for consumers), n (for purchase occasion), and w (for weekly purchase opportunities).

Brand Choice Elasticity

As in equation 1, the choice probability of brand j is given by

$$\text{(B1)} \qquad \mathbf{P}_j = \frac{\exp(\mathbf{b}'\mathbf{X}_j)}{\sum_{m=1}^{M} \exp(\mathbf{b}'\mathbf{X}_m)}.$$

If D_j is the deal variable for brand j, the brand choice elasticity with respect to D_j can be represented as

$$\text{(B2)} \qquad \eta_{BC(j)} = \frac{D_j}{p_j}\frac{\delta p_j}{\delta D_j} = bD_j(1 - p_j).$$

When interactions are included in the model, equation B2 is modified to

$$\eta_{BC(j)} = (b + \mathbf{b}^{*\prime}\mathbf{X}_j^*)\, D_j(1 - p_j)$$

where:

$\mathbf{X}_j^*$ = variables interacting with D_j and
$\mathbf{b}^*$ = parameters associated with the interaction terms.

Purchase Time Elasticity

The density function for interpurchase time is given by

$$f(t) = \alpha^2 t \exp(-\alpha t).$$

Therefore, the purchase incidence probability (p_t) for any unit time period (t = 1) is given by the cumulative density function,

$$\text{(B3)} \qquad p_t = 1 - (1 + \alpha)\exp(-\alpha),$$

where $\alpha = \exp(-\mathbf{c}'\mathbf{Y})$ and $\mathbf{Y}$ = market conditions (see equation 3).

Recall that the variables $\mathbf{Y}$ in the purchase time model are at the product category level. Thus the product category deal (Y_{deal}) is given as

$$Y_{\text{deal}} = \sum_{j=1}^{M} \delta_j D_j$$

where D_j = deal for brand j and δ_j = weight for brand j.

The model uses individual-level weights for δ_j. However, as indicated before, the elasticity analysis is being done at the average value of the market. We therefore use the brand's market share (s_j) as the weight.

If $\mathbf{Y} = \mathbf{Y}^* + Y_{\text{deal}}$,
c = parameter associated with Y_{deal}, and
$\mathbf{c}^*$ = parameters associated with $\mathbf{Y}^*$,

B3 can be written as

$$\text{(B4)} \qquad p_t = 1 - [1 + k\exp(-cs_jD_j)]\exp[-k\exp(-cs_jD_j)]$$

where $k = \exp\left(-\mathbf{c}^{*\prime}\mathbf{Y}^* - \sum_{m \neq j} cs_m D_m\right)$.

The purchase time elasticity therefore is given by

$$\text{(B5)} \qquad \eta_{PT} = \frac{D_j}{p_t}\frac{\delta p_t}{\delta D_j} = \frac{\alpha^2 cs_j D_j}{1 + \alpha - \exp(\alpha)}.$$

Purchase Quantity Elasticity

The expected purchase quantity is given by

$$Q = \sum_{k=1}^{L} kP(Q = k)$$

where $k = 1, 2, \ldots$ are discrete, ordered categories of purchase quantity. For example, in our case $k = 1$ implies 13 oz, $k = 2$ implies 16 oz, and so on. For representational clarity, we denote $k = 1$ as R_1, $k = 2$ as R_2, and so on. The expected purchase quantity therefore can be written as

$$\text{(B6)} \qquad Q = \sum_{c=1}^{L} R_c P_c$$

where $P_c = P(Q = c)$. From equation 8, we know that

$$\text{(B7)} \quad P_c = \frac{\exp(\theta_c - \boldsymbol{\beta}'\mathbf{Z})}{1 + \exp(\theta_c - \boldsymbol{\beta}'\mathbf{Z})} - \frac{\exp(\theta_{c-1} - \boldsymbol{\beta}'\mathbf{Z})}{1 + \exp(\theta_{c-1} - \boldsymbol{\beta}'\mathbf{Z})}.$$

As in the case of the purchase time model, here also the variables $\mathbf{Z}$ are at the product category level. Thus we can write

$$\boldsymbol{\beta}'\mathbf{Z} = \boldsymbol{\beta}^{*\prime}\mathbf{Z}^* + \beta s_j D_j$$

where:

- $\mathbf{Z}^*$ = all variables except deal of brand j with response parameters $\boldsymbol{\beta}^*$,
- D_j = deal of brand j with response parameter β, and
- s_j = market share of brand j.

Differentiating equation B7 with respect to D_j, we get

$$\eta_{PQ} = \frac{D_j}{Q}\frac{\delta Q}{\delta D_j} = \frac{D_j}{Q}\sum_{c=1}^{L} R_c \frac{\delta P_c}{\delta D_j}$$

$$= \frac{D_j}{Q}\sum_{c=1}^{L} R_c\, \beta\, s_j\, [P(Q \le R_{c-1})\, P(Q > R_{c-1}) - P(Q \le R_c)\, P(Q > R_c)].$$

REFERENCES

Ben-Akiva, Moshe E. and Steven R. Lerman (1985), *Discrete Choice Analysis: Theory and Applications to Travel Demand*. Cambridge, MA: MIT Press.

Blattberg, Robert C., Thomas Buesing, Peter Peacock, and Subrata K. Sen (1978), "Identifying the Deal Prone Segment," *Journal of Marketing Research*, 15 (August), 369–77.

———, Gary D. Eppen, and Joshua Lieberman (1981), "A Theoretical and Empirical Evaluation of Price Deals for Consumer Nondurables," *Journal of Marketing*, 45 (Winter), 116–29.

——— and Abel P. Jeuland (1981), "A Micro-Modeling Approach to Determine the Advertising Sales Relationship," *Management Science*, 27 (September), 988–1004.

Brown, Martin B. (1975), "The Asymptotic Standard Errors of Some Estimates of Uncertainty in the Two-Way Contingency Table," *Psychometrika*, 40 (September), 291–6.

Carpenter, Gregory S. and Donald R. Lehmann (1985), "A Model of Marketing Mix, Brand Switching, and Competition," *Journal of Marketing Research*, 22 (August), 318–29.

Chatfield, Christopher and Gerald J. Goodhardt (1973), "A Consumer Purchasing Model with Erlang Inter-Purchase Times," *Journal of the American Statistical Association*, 68 (December), 828–35.

Cooper, Lee G. and Masao Nakanishi (1988), *Market Share Analysis: Evaluating Competitive Marketing Effectiveness*. Boston: Kluwer Academic Publishers.

Doyle, Peter and John Saunders (1985), "The Lead Effect of Marketing Decisions," *Journal of Marketing Research*, 22 (February), 54–65.

Dunn, R., S. Reader, and N. Wrigley (1983), "An Investigation of the Assumptions of the NBD Model as Applied to Purchasing at Individual Stores," *Applied Statistics*, 32(3), 249–59.

Ehrenberg, Andrew S. C. (1959), "The Pattern of Consumer Purchases," *Applied Statistics*, 8 (March), 26–41.

——— (1972), *Repeat Buying: Theory and Applications*. Amsterdam: North Holland Publishing Company.

Gensch, Dennis H. and Wilfred W. Recker (1979), "The Multinomial, Multiattribute Logit Choice Model," *Journal of Marketing Research*, 16 (February), 124–32.

Goodman, Leo A. and William H. Kruskal (1954), "Measures of Association for Cross Classification," *Journal of the American Statistical Association*, 49 (December), 732–64.

Guadagni, Peter M. (1983), "A Nested Logit Model of Product Choice and Purchase Incidence," in *Proceedings of the 1983 ORSA/TIMS Marketing Science Conference*, Fred S. Zufryden, ed. Los Angeles: ORSA/TIMS, 90–100.

——— and John D. C. Little (1983), "A Logit Model of Brand Choice Calibrated on Scanner Data," *Marketing Science*, 2 (Summer), 203–38.

Hauser, John R. (1978), "Testing the Accuracy, Usefulness and Significance of Probabilistic Choice Models: An Information Theoretic Approach," *Operations Research*, 26 (May), 406–21.

Herniter, Jerome (1971), "A Probabilistic Market Model of Purchase Timing and Brand Selection," *Management Science*, 18 (December), Part II, P102–13.

Jones, Morgan and Fred Zufryden (1980), "Adding Explanatory Variables to a Consumer Purchase Behavior Model: An Exploratory Study," *Journal of Marketing Research*, 17 (August), 323–34.

Keon, John (1980), "The Bargain Value Model and a Comparison of Managerial Implications with the Linear Learning Model," *Management Science*, 26 (November), 1117–30.

Kinberg, Yoram, Ambar G. Rao, and Melvin F. Shakun (1974), "A Mathematical Model for Price Promotions," *Management Science*, 20 (February), 948–59.

Krishnamurthi, Lakshman and S. P. Raj (1985), "The Effect of Advertising on Consumer Price Sensitivity," *Journal of Marketing Research*, 22 (May), 119–29.

——— and ——— (1988), "A Model of Brand Choice and Purchase Quantity Price Sensitivities," *Marketing Science*, 7 (Winter), 1–20.

Kuehn, Alfred A. and Albert C. Rohloff (1967), "Consumer Response to Promotions," in *Promotional Decisions Using Mathematical Models*, Patrick J. Robinson, ed. Boston: Allyn & Bacon, 45–145.

Kunreuther, Howard (1973), "Why the Poor May Pay More for Food: Theoretical and Empirical Evidence," *Journal of*

Business, 46 (July), 368–83.
Lattin, James M. (1987), "A Model of Balanced Choice Behavior," *Marketing Science*, 6 (Winter), 48–65.
Lilien, Gary L. (1974), "A Modified Linear Learning Model of Buyer Behavior," *Management Science*, 20 (March), 1027–36.
Little, John D. C. and Peter Guadagni (1986), "When and What to Buy: A Hierarchical Model of Coffee Purchase," paper presented at the 1986 ORSA/TIMS Conference, Dallas, TX.
Massy, William F. and Ronald E. Frank (1965), "Short Term Price and Dealing Effects in Selected Market Segments," *Journal of Marketing Research*, 2 (May), 171–85.
———, David Montgomery, and Donald G. Morrison (1970), *Stochastic Models of Buying Behavior*. Cambridge, MA: MIT Press.
McCullagh, Peter (1980), "Regression Models for Ordinal Data," *Journal of the Royal Statistical Society*, 42, Series B, 109–42.
McFadden, Daniel (1974), "Conditional Logit Analysis of Qualitative Choice Behavior," *Frontiers in Econometrics*, Paul Zarembka, ed. New York: Academic Press, Inc.
Morrison, Donald G. and David C. Schmittlein (1981), "Predicting Future Random Events Based on Past Performance," *Management Science*, 27 (September), 1006–23.
Naert, Philippe A. and Peter S. Leeflang (1978), *Building Implementable Marketing Models*. Boston: Leiden Press.
——— and M. Weverbergh (1981), "On the Prediction Power of Market Share Attraction Models," *Journal of Marketing Research*, 28 (May), 146–53.
Neslin, Scott A., Caroline Henderson, and John Quelch (1985), "Consumer Promotions and the Acceleration of Product Purchases," *Marketing Science*, 4 (Spring), 147–65.
——— and Robert W. Shoemaker (1983), "Using a Natural Experiment to Estimate Price Elasticity: The 1974 Sugar Shortage and Ready-to-Eat Cereal Market," *Journal of Marketing*, 47 (Winter), 44–57.
Ortmeyer, Gwen K. (1985), "A Model of Consumer Response to Promotions," unpublished dissertation, Stanford University.
Paull, Allen E. (1978), "A Generalized Compound Poisson Model for Consumer Purchase Panel Data Analysis," *Journal of the American Statistical Association*, 73 (December), 706–13.
Punj, Girish N. and Richard Staelin (1978), "The Choice Process for Graduate Business Schools," *Journal of Marketing Research*, 15 (November), 588–98.
Raj, S. P., Richard Staelin, and Andrew A. Mitchell (1977), "A General Procedure for Measuring the Effect of Promotions on Buyer Behavior," *American Statistical Association Proceedings*, 112–16.
Shoemaker, Robert W. (1979), "An Analysis of Consumer Reactions to Product Promotions," *American Marketing Association Proceedings*, 244–48.
Theil, Henri (1965), *Economic Forecasts and Policy*. Amsterdam: North Holland Publishing Company.
Wagner, Udo and Alfred Taudes (1986), "A Multivariate Polya Model of Brand Choice and Purchase Incidence," *Marketing Science*, 5 (Summer), 219–44.
Ward, Ronald W. and James E. Davis (1978), "A Pooled Cross-Section Time Series Model of Coupon Promotions," *American Journal of Agricultural Economics*, 60 (August), 393–401.
Wheat, Rita and Donald G. Morrison (1986), "Exploratory Data Analysis Applied to Scanner Data: How Regularly Do Consumers Purchase—and Does it Matter," paper presented at the 1986 ORSA/TIMS Conference, Dallas, TX.
Wilson, R. D., L. M. Newman, and Manoj Hastak (1979), "On the Validity of Research Methods in Consumer Dealing Activity: An Analysis of Timing Issues," in *1979 AMA Educators' Proceedings*, Neil Beckwith et al., eds. Chicago: American Marketing Association, 41–6.

Reprint No. JMR254101

After July 1, 1988, please submit manuscripts to Michael J. Houston, Editor, Carlson School of Management, University of Minnesota, 271 19th Avenue South, Minneapolis, MN 55455.