A CHOICE MODELING APPROACH TO EVALUATE EFFECTIVENESS OF BRAND DEVELOPMENT INITIATIVES

ARINDAM BANERJEE
Associate Professor, Marketing Area
Indian Institute of Management Ahmedabad
Vastrapur, Ahmedabad
Gujarat, Pin 380015
India
Email: arindam@iimahd.ernet.in
Tel: 91-79-26324898

DHEERAJ AWASTHY
Doctoral candidate, Fellow Program in Management
Indian Institute of Management Ahmedabad
India
Email: dheeraj@iimahd.ernet.in

VIVEK GUPTA
Doctoral candidate, Fellow Program in Management
Indian Institute of Management Ahmedabad
India
Email: vgupta@iimahd.ernet.in
ABSTRACT

We describe the application of a nested logit function for modeling brand choice using household transaction data from the Indian market. This is unique since it is one of the first attempts to integrate disparate consumer information sources available at various levels of aggregation towards developing a prediction model for brand market share. We further develop a methodology for brand market share decomposition into components that can be attributable to various explanatory variables. The implications are significant since this methodology helps in using behavioral tracking data towards developing a decision tool to evaluate marketing programs.

KEYWORDS: BRAND DEVELOPMENT, PREDICTION MODEL, CHOICE MODELING, NESTED LOGIT, EMERGING MARKETS, MARKETING DSS
A CHOICE MODELING APPROACH TO EVALUATE EFFECTIVENESS OF BRAND DEVELOPMENT INITIATIVES

1 INTRODUCTION

Effective decision making requires ongoing precise feedback on performance, which identifies good vs. bad decisions taken in the past. If managers could intelligently use such a feedback system to direct precious marketing resources towards initiatives that are more efficient, managerial decisions would be governed by a significant element of scientific rationality rather than subjective heuristics. Predictive models with sufficient robustness can provide such effective feedback systems that may translate into competitive advantage to firms employing these tools.

Prediction models can be used to map input variables like sales promotions, price variations, distribution, and advertising to defined output performance variables like market share, returns etc. These models provide a logical basis to the manager to compute the differential impact of firm’s marketing strategies vis-à-vis its competitors on the market share of its brands using scenario builders. These models could be based on volumetric share analysis, choice share analysis or a combination of the two. Of these, volumetric share based models are very popular in the west because of low complexity and easy interpretability.

Development of such models in emerging markets has not been attempted in the past because of non-availability of large-scale databases, which are necessary to build robust models. In the recent past, many organizations have invested in fairly large sized panels to track customer purchases in India. However, most of the applications based on these databases have been limited to tracking customer purchases on an ongoing basis. There has been little attempt at developing behavior models that help identify drivers of consumer behavior and their relative impact. With the current data gathering practices easier volumetric modeling is not feasible.

Our paper is perhaps the first significant attempt to utilize customer transaction data from a large emerging market to build a prediction model based on choice share analysis. We also develop a decomposition technique for linking performance to various marketing mix elements from consumer choice models. The model has high managerial relevance because it directly links marketing initiatives to performance
outcomes such as competitive market-shares. The paper also describes the challenges of estimating models by integrating disparate customer databases with varying quality of data.

The paper is divided in five sections. First section provides the introduction, second section discusses the rising importance and problems of emerging markets, third section discusses about choice of modeling technique in such markets and fourth section discusses model development and results. Conclusions are provided in fifth section.

2 MODELLING ISSUES IN EMERGING MARKETS

Emerging markets term is used to characterize developing markets which have higher growth rates, increasing incomes and are supposedly the future drivers of world economy in 21st century. These markets are different from their developed counterparts in terms of their market characteristics and data gathering practices. They are more volatile and constantly evolving. For example, the number of SKU’s in India in the consumer retail sector has grown from 9093 to 17739 in the period from 1990 to 1996 (Banerjee, Raghuram and Koshy, 1999). They have low penetration of organized retailing outlets and a significantly large concentration of small convenience stores with low or negligible automation.

The heterogeneity in modes of selling, low automation create various problems in data capture and collation (Banerjee and Banerjee, 2000). Overall market is organized in haphazard manner with different agencies capturing separate data using different attributes; hence compatibility of databases is a significant problem. Much of this information is gathered for tracking overall market trends and hence is collected at the aggregated level.

A major source of consumer level data capture in these markets is through the syndicated panels maintained by companies themselves. The household level syndicated panel information in these markets is still gathered using diary mode. The rich transaction level information, which is easily available in scanner databases in other developed markets, is not available in these databases. While it is essential to improve data collection practices, it would take a considerable amount of time to evolve a system that is comparable to the west. Hence, a modeling methodology that is compatible to existing data-gathering practices can have much higher managerial utility in near future.
3 PREDICTION MODELS: CHOICE OF AN UNDERLYING TECHNIQUE

Prediction models using consumer behavior data normally link marketing mix variables to output performance measured as brand choice or quantity purchased. These measures at a higher level of aggregation can be considered as good surrogates for the brand’s market share. The models can be estimated either using store level data gathered from sample of stores or transaction level data collected from a sample of consumers. We preferred to estimate a brand choice model (McFadden, 1986) since the nature of the product (Toiletries – large packs) largely constrained a large percentage of the sample respondents to buy single packs on any purchase occasion. Hence, choice of brand purchased was a more pertinent to study than purchase quantity.

Consumer choice model uses utility maximization theory as its foundation. It models the choice of a consumer of a particular brand at a particular time as generating maximum utility given attributes of available alternatives and consumer’s demographic and psychographic characteristics. Consumer choice has been modeled in literature as brand choice or simultaneous decision of both choice and quantity (Chiang, 1991, Guadagni and Little, 1983). Logit and probit are most commonly used specifications in the literature to model the phenomenon (Guadagni and Little, 1983, Kamakura and Srivastava, 1984, Moore and Lehmann, 1989, Raju, Dhar and Morrison, 1994). Consumer choice models are developed using data pertaining to social, demographic, psychometric measures of customers and marketing mix variables of all competing products in a product category gathered over large number of choice occasions (McFadden, 1986).

The required data can be obtained through experimental studies (behavioral intention data) or scanner data sources (revealed preference data) or field surveys (either behavior intention/ revealed preference data) (refer figure 1).

---------------------------------------------------------------
Insert Figure 1 around here
---------------------------------------------------------------

Experimental studies are a rich source of data where the researcher can design a set up to cover all the levels of a limited number of critical attributes required for choice modeling (McFadden, 1986). Though
experimental studies provide rich information, the external validity of experimental data has been a matter of debate in the literature (Thye, 2000). In contestable markets with continuous entry and exit of brands, there is a continuous shift in underlying structure of the market thus requiring refitting of model on any major entry/exit. For instance, there were 57 core categories of products in 1990 in India which grew to 76 by 1996 (Banerjee, Raghuram and Koshy, 1999). Hence, calibrating and recalibrating models using data from experimental setups may prove to be very costly.

Scanner data collected from syndicated sources has been used at three levels of aggregation for choice modeling in the western context – household level, store level, and market level (Gupta and Chintagunta, 1996). Data is not available in similar forms in the emerging markets. Instead substitutes are available in the form of transaction panel, behavior panel data as discussed above.

Field surveys have been used in the literature to capture attitude and perception data (Ben-Akiva, et al., 1999) which is further merged with data from other sources to estimate prediction models. This is however a fairly complicated process.

Given the constraints of available data and the hypothesis on the nature of purchase behavior in the specific product category used for study, it was concluded that estimating a choice model using transaction data from a consumer panel along with relevant augmentation from other secondary sources of data, was the most appropriate prediction model for the context.

4 MODEL DEVELOPMENT

4.1 DATASETS USED IN THE STUDY

The data used for the study was made available to the researchers by a leading firm operating in consumer packaged good market in India. The data provided was from Oct 99 to Dec 2000. It comprised three distinct datasets.

**Purchase transaction (behavior) data**

The firm maintained a household diary panel with the help of a professional market research agency to collect transaction level information from the randomly selected households in a single city. The data was
recorded manually in a dairy by the consumer soon after his/her purchase and was collected by the agency on a monthly basis. Data was collected from each panelist for brands bought, size (SKU) bought, amount consumed, frequency of purchase and demographic variables. It did not record point of purchase marketing mix elements. Main features of data are given in table 1.

\[
\begin{array}{|c|}
\hline
\text{Insert Table 1 around here} \\
\hline
\end{array}
\]

Attitudinal data
Firm maintained a revolving panel at weekly level of 50 consumers to track brand related image/attitude association with the help of a professional market research agency. The data was collected for all competing brands.

Marketing Mix Variables
Information about aggregate level marketing mix variables was available from retail audit data of ORG-MARG\(^1\) which captures the information about pricing and distribution of all the brands in a category for a sample of stores in each market on a monthly basis. This data was available at aggregate level for the market and not at the transaction level. Additionally, data was available on the number of ‘promoted brands’ sold in each month but there was no information available on the type of promotion strategy used.

4.2 CHOICE MODEL

This section describes the development of choice model, aggregation to compute a surrogate market share and decomposition of market share to components attributable to marketing mix variables.

We derived the brand choice model using a logistic function based on McFadden (1986). The general form of model can be described as:

\[
\text{Brand Choice} = f (\text{Marketing Mix, Psychometric Variables, Demographic variables}) \rightarrow (1)
\]

\(^1\) a leading market research company in India
In literature, marketing mix variables like regular price, promoted price, display, type of promotion have been used to model choice (Guadagni and Little, 1983). We also include attitude variables as explicit treatment of psychological factors in choice models leads to a more behaviorally realistic representation of the choice process, and consequently, better explanatory power (Ben-Akiva and Boccara, 1995, Ben-Akiva et al., 1999, Kalidas, Dillon and Yuan, 2002, McFadden, 1986). We use demographic variables in choice models as it provides managers with identifiable segments and their choice priorities (Gupta and Chintagunta, 1994, Kalyanam and Putler, 1997). We incorporated as many relevant market and consumer specific variables accessible to the managers to allow for the development of a robust yet useful model.

We use equation (1) to develop our choice model including marketing mix (refer table 2), demographics (refer table 3) and attitude variables. We model consumer choice as choice of brand and pack-size only with number of packs purchased being one at any choice occasion. The quantity decision was not considered because mean packs purchased per choice occasion was marginally over one (1.2). In this model we derive the computation of choice utility as given below.

Let us consider that an individual ‘j’ is faced with a situation in which s/he has to choose amongst ‘m’ brands available in market. Let ‘P_{jk}’ is the probability that individual ‘j’ will choose brand ‘k’ from the choice set ‘m’; ‘X_j’ is the demographic characteristics of the individual and ‘Z_{jk}’ is the characteristic of ‘k^{th}’ brand as observed by individual ‘j’ (brand related factors like price, distribution etc.) \( \alpha \) and \( \beta \) are parameter estimates for conditional and polytomous variables.

\[ P_{jk} = \frac{\exp(\alpha Z_{jk} + \beta X_j)}{\sum_{l=1}^{m} \exp(\alpha Z_{jl} + \beta X_l)} \quad \rightarrow (2) \]

The model was developed for the premium laundry detergent category using the data which was available from October 1999 to December 2000 from the Mumbai market in India. There were 930 households in the dataset with 4165 transactions. The overall choice set in December 2000 constituted four detergent
brands – Alpha, Beta, Gamma and Theta\textsuperscript{2}. Theta was introduced in the market in June 2000. Alpha and Beta constituted around 85% of market and thus were major brands.

To incorporate the brand and pack-size choice a two stage model (nested logit) was estimated (see Figure 2). In stage 1, the utility because of pack sizes was computed and in stage 2, pack coefficient estimated in stage 1 were used as an additional variable and probability of choice of each brand was estimated (Maddala 1986).

Drawing from Guadagni and Little (1983) we have computed the monthly market shares of brands as average probability of choice of brand in a month over all choice occasions. This aggregated market share was decomposed to various marketing mix elements. The decomposition procedure is discussed in a later section.

4.3 DISCUSSION ON VARIABLES: DEVELOPING PROXIES

4.3.1 Marketing Mix Variables

For the current study, data on marketing mix variables (price, distribution) was available at the market level from the retail audit data maintained by ORG-MARG. We operationalized price as Maximum Retail Price per gram in each month. Price discounting is not a frequently employed marketing tactic and hence MRP is the only operational price variable. All consumers face similar prices as indicated in retail survey even for different transaction in each month. The use of coupons as marketing tool in emerging markets is very low, as compared to developed markets.

We used aggregate level dealer stocking position (% of dealers stocking the brand) as a proxy for availability of the brand at transaction level. We operationalized promotion intensity as a continuous variable measuring the proportion of promoted brand sold in a month to overall promotion brands sold during the total period in the transaction panel data.

Summary of the operationalized marketing mix variables used to estimate the model is provided in table 2.

\textsuperscript{2} Brand names have been camouflaged to maintain confidentiality.
4.3.2 Psychometric Variables

We have modeled these variables as consumer attitude towards various brands in the choice set. We believe this is superior and more appropriate as compared to use of advertising budgets because (1) attitude have a more direct relationship with choice than advertising budgets (2) Inclusion of advertisement budges assumes that all brand advertisements are equally effective, which may not be true. Also brand level advertising budgets are not available in India.

The attitude data was available as attitude scores were measured on 18 statements on dichromatic (yes/no) scale. However these attitudinal statements were highly correlated. On applying factor analysis (after Varimax Rotation) the 18 attitude statements converged into three distinct factors. We incorporated three statements (variables) which had highest loading on these factors dropping all other statements.

4.3.3 Demographic Variables

Demographic variables were available for each household for head of household and respondent. We applied a classification tree analysis using brand choice as independent variable and demographics as dependent on CART (Vs 4.0) to create segments with interaction of demographic variables. These interactions were coded as dummy variables and included in the model. This was done to capture significant interactions within demographic variables.

A schematic representation of cart dummies is shown in figure 3.

The operationalization of demographic variables is given in table 3.
4.4 MODEL ESTIMATION

The model was estimated using PHREG procedure of SAS 8.01. The likelihood function was customized to model three alternatives (brands) from Oct-99 to May-00 and four alternatives (brands) from June-00 to Dec-00.

4.5 QUALITY OF FIT

Model Fit was comparable with typical choice estimation in research literature. The pseudo R² (similar to Maddala (1986)) of the estimated model was 0.49. Fit results reported are at individual transaction level and not at the market aggregate level. Hit rate was computed to obtain predictive validity of the model. The hit-rate of the model was 70.23%, i.e. in 70.23% of records in our sample the model predicted choice matched with actual choice of consumer. The model was also validated on data provided for Jan 2001- Jun 2001 period.

Aggregate market shares (similar to (Guadagni and Little,1983)) were computed at monthly, quarterly and semi-annually level using estimated coefficients and was compared to actual market share of the panel data. The charts comparing actual market shares to predicted market shares for both calibration sample and validation sample are given in figure 4 & 5. In calibration sample the absolute standard deviation in the predicted and actual market share of alpha brand was 3.47 and beta brand was 2.77, which shows that the goodness of fit of model was within acceptable limits. In validation sample the absolute standard deviation in predicted and actual share was 3.06 for brand alpha and 4.5 for brand beta.

4.6 DISCUSSION OF COEFFICIENTS

The impact coefficients of significant (at 90% level) marketing mix, attitude variables (conditional variables) and polytomous variables are reported in table 4.
The two composite attitudinal variables Economy and Performance came significant and positive in direction. The pack level composite variable based derived from stage one logit model was also significant, thereby meaning that the pack level tactics had a significant effect on overall brand choice. The parameter estimate of distribution was negative and possibly it was because of aggregate nature of marketing mix variables. The other possible reason for such results can be that one of the major brands i.e. Alpha in the corresponding period reduced its distribution and applied a more focused distribution strategy. The CART based structural dummy variables and many polytomous demographic variables also were significant in the results. This indicates to an underlying preference for a particular brand among a class of consumers.

4.7 MANAGERIAL IMPLICATIONS: DEVELOPING A FRAMEWORK TO EVALUATE MARKETING SPENDS

The model estimated in the previous section has a multiplicative form. The broad elements of the model that influence the final market share of the brand are:

\[
\text{Market Share} = \text{Base Share}^3 \times \text{Effect due to Advertising effort} \times \text{Effect due to sales effort} \rightarrow (3)
\]

The parameters (beta coefficients) estimated from the model, although represent the impacts of their specific mix variable, are nevertheless unintuitive to most practicing managers. A better indication of the impact of the various marketing mix elements would be the contribution of total market share attributable to each variable. Given the nature of the data available, and the current managerial imperatives, it was necessary to develop an algorithm to separate out the composite effect of advertising and base brand development initiatives from other field level initiatives. For this purpose the market share was decomposed into two constituents (1) Composite market share due to Baseline brand equity and advertising (MSA) (2) Market share due to other factors (MSO).

3 Base share is the share of firm’s brand because of brand equity or long term sales/advertising effort.
The objective of this decomposition exercise is simply to devise a suitable index for measuring the health of a brand, an issue that confronts most brand managers in a competitive environment. A higher percentage of MSA in the total share would imply lower vulnerability to competitive pressures due to stronger core equity of the brand. At a more tactical level, the proportion of MSA and MSO is indicative of the performance of brand management relative to the sales function in maintaining the health of the brand.

The decomposition algorithm computed the effect of one variable (attitudes based on advertising) by calculating market share based on the logit model controlling for the impact of all other variables. With the estimated beta coefficients from the choice model, the probabilities of choice at individual transaction level were computed by replacing the marketing mix variables with their modal values for all variables except for the variables measuring brand attitudes. The demographic variables in the model were replaced with the average value for each demographic variable computed from the sample. This represented the base utility for each brand as perceived by the demographic profile of the sample.

Let $D_k$ be the average utility due to the sample demographic profile for brand $k$. This is computed by summing up the actual value of every relevant demographic variable in the model weighted by their beta coefficients. This operation is done for every transaction record in the data and then averaged by brand across all records.

Let $A_{jk}$ be the utility for brand $k$ due to brand attitudes during transaction $j$.

Let $M_k$ be the utility for brand $k$ due to all marketing mix variables except attitudes across all transactions at the modal value obtained from the sample.

Market-Share due to Advertising for brand $k$ (MSA) = $D_k * A_{jk} / \sum_{all \ brands} (D_k * M_k * A_{jk})$

Market share due to Other Factors (MSO) = Actual brand market share – MSA.

While there can be competing methodologies for decomposing the market share of the brand, the approach outlined above seemed to be the most palpable to a group of practicing managers whom we presented the findings. Although the specific methodology used will drive the nature of results and hence it is necessary to
be logically consistent, the highlight of this exposition is the development of a framework using customer
data to diagnose brand health and directly relate it to managerial initiatives. The two decomposed
components are plotted for each brand by month over a 13-month period as shown in Figure 6 & 7 (for Alpha
and Beta).

The charts show that brand alpha has strong core since MSA is very significant. In fact, the core equity is
not supported well by “field selling” initiatives (other initiatives) because the actual market share is lower
than the MSA. Brand “Beta” seems to have on the whole a lower market share, which is also reflected in a
weaker core (MSA). However, this brand seems to be ably supported in the markets through “field sales
management” initiatives (MSO). Qualitative findings based on market perceptions and reactions from
managers of a large consumer marketing company validated these claims.

5 CONCLUSION
In this paper we present a novel combination of tools and techniques to develop a predictive model for
emerging markets linking marketing mix variables to market share performance. The model development
and decomposition methodology can help managers to evaluate the performance of the initiatives like brand
building programs or promotion programs. It can also assist managers to apportion the limited marketing
resources for meaningful investment to long term market building activities.

While it is difficult to significantly modify investments made in information gathering in the near term,
proper tooling to suit the sub optimal data quality can provide significant gains to managers in building
decision support infrastructure. We provide a way to exploit large volumes of available data with the
companies for developing appropriate decision support systems for managers in emerging markets.

Due to limitations in the existing data gathering practices, the model does not incorporate various relevant
marketing mix variables. For instance, qualitative findings indicate the importance of distribution variable
(availability of brand at POP) in the Indian markets. However, the pertinent information is available only at
the market aggregate level and hence its ability to explain transaction level choice behavior is limited. It
highlights the need for a planned investment in customer data in the emerging markets for managers to derive maximum benefits of prediction based models.
REFERENCES


Maddala GS. Limited Dependent and Qualitative Variables in Econometrics: Cambridge University Press, 1986.


**TABLE 1 - CHARACTERISTIC OF BEHAVIOR PANEL DATA**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Number of Households</td>
<td>930</td>
</tr>
<tr>
<td>Total Number of Transactions</td>
<td>4165</td>
</tr>
<tr>
<td>Transactions of Brand Alpha</td>
<td>1987</td>
</tr>
<tr>
<td>Transactions of Brand Beta</td>
<td>1689</td>
</tr>
<tr>
<td>Transactions of Brand Gamma</td>
<td>429</td>
</tr>
<tr>
<td>Transactions of Brand Theta</td>
<td>60</td>
</tr>
<tr>
<td>Maximum no. of brands purchased by any customer in one month</td>
<td>12</td>
</tr>
<tr>
<td>Average no. of brands Purchased per month</td>
<td>1.013</td>
</tr>
<tr>
<td>Average no. of brands Purchased in two years (surrogate for loyalty)</td>
<td>1.4</td>
</tr>
<tr>
<td>Average number of transactions made in two years</td>
<td>4.469</td>
</tr>
<tr>
<td>Average number of packs bought per transaction</td>
<td>1.235</td>
</tr>
</tbody>
</table>
### TABLE 2 – MARKETING MIX VARIABLES

<table>
<thead>
<tr>
<th>Name</th>
<th>Unit</th>
<th>Type</th>
<th>Collection Level</th>
<th>Data Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price</td>
<td>Rs per gm</td>
<td>Ratio</td>
<td>Monthly market level</td>
<td>Retail</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Scaled</td>
<td>aggregate</td>
<td>Survey</td>
</tr>
<tr>
<td>Promotion</td>
<td>%promotion</td>
<td>Ratio</td>
<td>Monthly market level</td>
<td>Retail</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Scaled</td>
<td>aggregate</td>
<td>Survey</td>
</tr>
<tr>
<td>Distribution</td>
<td>% Dealer</td>
<td>Ratio</td>
<td>Monthly market level</td>
<td>Retail</td>
</tr>
<tr>
<td></td>
<td>stocking</td>
<td>Scaled</td>
<td>aggregate</td>
<td>Survey</td>
</tr>
</tbody>
</table>
# TABLE 3 DEMOGRAPHIC VARIABLES

<table>
<thead>
<tr>
<th>Name</th>
<th>Unit</th>
<th>Type</th>
<th>Level</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age of head of household</td>
<td>Dummy</td>
<td>Nominal (12 categories)</td>
<td>Disaggregate</td>
<td>Household panel</td>
</tr>
<tr>
<td>Age of Respondent</td>
<td>Dummy</td>
<td>Nominal (12 categories)</td>
<td>Disaggregate</td>
<td>Household panel</td>
</tr>
<tr>
<td>Education of Head of household</td>
<td>Dummy</td>
<td>Nominal (9 Categories)</td>
<td>Disaggregate</td>
<td>Household panel</td>
</tr>
<tr>
<td>Education Respondent</td>
<td>Dummy</td>
<td>Nominal (9 Categories)</td>
<td>Disaggregate</td>
<td>Household panel</td>
</tr>
<tr>
<td>Respondents Knowledge of English</td>
<td>Dummy</td>
<td>Nominal (6 categories)</td>
<td>Disaggregate</td>
<td>Household panel</td>
</tr>
<tr>
<td>Medium of Education of Head of Household</td>
<td>Dummy</td>
<td>Nominal (5 categories)</td>
<td>Disaggregate</td>
<td>Household panel</td>
</tr>
<tr>
<td>Medium of Education of respondent</td>
<td>Dummy</td>
<td>Nominal (5 categories)</td>
<td>Disaggregate</td>
<td>Household panel</td>
</tr>
<tr>
<td>Marital status respondent</td>
<td>Dummy</td>
<td>Nominal (5 categories)</td>
<td>Disaggregate</td>
<td>Household panel</td>
</tr>
<tr>
<td>Occupation head household</td>
<td>Dummy</td>
<td>Nominal (15 categories)</td>
<td>Disaggregate</td>
<td>Household panel</td>
</tr>
<tr>
<td>Occupation respondent</td>
<td>Dummy</td>
<td>Nominal (15 categories)</td>
<td>Disaggregate</td>
<td>Household panel</td>
</tr>
<tr>
<td>Household income</td>
<td>Rs. per family member</td>
<td>Ratio Scaled</td>
<td>Disaggregate</td>
<td>Household panel</td>
</tr>
<tr>
<td>Mother Tongue</td>
<td>Dummy</td>
<td>Nominal (19 categories)</td>
<td>Disaggregate</td>
<td>Household panel</td>
</tr>
<tr>
<td>Working Status</td>
<td>Dummy</td>
<td>Nominal (5 categories)</td>
<td>Disaggregate</td>
<td>Household panel</td>
</tr>
<tr>
<td>Owner of Washing Machines</td>
<td>Dummy</td>
<td>Nominal (2 categories)</td>
<td>Disaggregate</td>
<td>Household panel</td>
</tr>
<tr>
<td>Segment Dummy</td>
<td>Dummy</td>
<td>Nominal(2 categories)</td>
<td>Disaggregate</td>
<td>CART</td>
</tr>
<tr>
<td>Variable</td>
<td>Parameter Estimate</td>
<td>Significance Level</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>--------------------</td>
<td>--------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Economy</td>
<td>0.87</td>
<td>0.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Performance</td>
<td>0.55</td>
<td>0.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distribution</td>
<td>-0.02</td>
<td>0.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PACK Composite</td>
<td>0.29</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Awareness</td>
<td>0.03</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CART1_Beta</td>
<td>-0.70</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CART1_Gamma</td>
<td>-0.80</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CART1_Theta</td>
<td>-1.70</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CART2_Beta</td>
<td>0.87</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CART4_Beta</td>
<td>-0.16</td>
<td>0.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CART4_Theta</td>
<td>-0.94</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CART6_Gamma</td>
<td>-1.04</td>
<td>0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CART6_Theta</td>
<td>-2.03</td>
<td>0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CART7_Gamma</td>
<td>1.26</td>
<td>0.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beta_Repondent Speaks &amp; Reads English</td>
<td>-0.28</td>
<td>0.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beta_Repondent Speaks English</td>
<td>0.44</td>
<td>0.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beta_Medium of Education_Hindi</td>
<td>0.21</td>
<td>0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beta_Married Housewife</td>
<td>-0.54</td>
<td>0.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beta_Occupation_Business_0Employee</td>
<td>0.29</td>
<td>0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beta_Occupation_Other</td>
<td>0.45</td>
<td>0.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beta_washing machine owner</td>
<td>0.25</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beta_Full time working</td>
<td>-0.39</td>
<td>0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beta_Part time working</td>
<td>-0.28</td>
<td>0.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gamma_Age of Head Household 41-44 YEARS AGE</td>
<td>0.69</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gamma_Age of Head Household 59 + YEARS AGE</td>
<td>0.26</td>
<td>0.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gamma_Age of Respondent_Others</td>
<td>-0.96</td>
<td>0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gamma_Medium of education_other</td>
<td>0.54</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gamma_Mother Tongue_Bengali</td>
<td>0.82</td>
<td>0.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gamma_Mother Tongue_Malyalam</td>
<td>1.59</td>
<td>0.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gamma_Mother Tongue_Others</td>
<td>-0.58</td>
<td>0.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gamma_Mother Tongue_Sindhi</td>
<td>-0.77</td>
<td>0.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gamma_Married Housewife</td>
<td>0.85</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gamma_Occ_Business_1-9 Employee</td>
<td>-2.05</td>
<td>0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gamma_Occupation_Other</td>
<td>-1.02</td>
<td>0.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gamma_Occupation_Self Employed</td>
<td>-0.62</td>
<td>0.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gamma_Occupation_Shop</td>
<td>-0.67</td>
<td>0.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gamma_Occ_Unskilled Worker</td>
<td>-0.66</td>
<td>0.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gamma_Full Time Working</td>
<td>-0.96</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gamma_Part Time Working</td>
<td>-0.74</td>
<td>0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Theta_Age of Respondent_Others</td>
<td>-1.73</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Theta_Occupation_Unskilled</td>
<td>0.95</td>
<td>0.01</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
FIGURE 1 – DATA REQUIREMENTS FOR CHOICE MODELS

- **Marketing Mix**
  - Price
  - Promotion
  - Distribution

- **Demographics**
  - Income
  - Age
  - Marital Status
  - Education (etc.)

- **Psychometric Variables**
  - Attitude

- **Consumer Choice**
FIGURE 2 - STRUCTURE OF MARKET

Stage 2
Brand Choice

Alpha
Large Pack
Medium Pack
Small Pack

Theta
Small Pack

Gamma
Large Pack

Beta
Large Pack
Medium Pack
Small Pack

Stage 1
FIGURE 3 - CART STRUCTURAL DUMMIES

Split based on Age of Respondent

Split Based on Occupation of Head of Household

Split based on Age of Household

Beta/Gamma 15%

Age: All categories

Beta - 34%

Age: Middle aged

Alpha - 3%

Occupation: Except Officers/Middle Senior

Alpha - 8%

Age: 20-30 yrs.

Segment’s Preference, Description & Size Given
FIGURE 4 - BRAND ALPHA MONTHLY AGGREGATED MARKET SHARE
FIGURE 5 - BRAND BETA MONTHLY AGGREGATED MARKET SHARE
FIGURE 6 - MARKET SHARE DUE TO (ADVERTISING + BASELINE) AND ACTUAL MARKET SHARE FOR ALPHA BRAND
FIGURE 7 - MARKET SHARE DUE TO (ADVERTISING + BASELINE) AND ACTUAL MARKET SHARE FOR BETA BRAND