MONASH

WP 17/96

ISSN 1032-3813
ISBN 0 7326 0798 1

MONASH UNIVERSITY

ANCORA IMPARO

AUSTRALIA

A LOGIT MODEL OF LAUNDRY DETERGENT BRAND CHOICE IN MELBOURNE

Tim R.L. Fry and Ritchard Longmire

Working Paper 17/96
September 1996

DEPARTMENT OF ECONOMETRICS

A Logit model of laundry detergent brand choice in Melbourne*.

Tim R.L. Fry and Ritchard Longmire
Department of Econometrics
Monash University
Clayton, Vic. 3168
Australia

Email: Tim.Fry@BusEco.monash.edu.au

September 1996

Abstract: The Consumer Panel of Australia data collected by The Roy Morgan Research Centre is used to investigate the determinants of brand choice behavior in the laundry detergent market. To reduce potential problems of heterogeneity in the market all purchases of the eleven leading brands in the Melbourne metropolitan area over the period July 1992 to June 1993 are considered. The purchases over this period are divided into an estimation sample and a holdout sample. A mixed Logit model is then fitted using the estimation sample. Price, brand loyalty, household size and household income are found to be significant in explaining brand choice behavior. A test for the, potentially restrictive, property of Independence of Irrelevant Alternatives (IIA) is carried out and it is found that IIA is acceptable for this data. The model appears to fit well and estimated elasticities are presented. The estimated model is then validated using the holdout sample. It is found that the model is very good at predicting the brand choices made by consumers in the holdout sample.

JEL Classification: M31, C25

Keywords: Logit, brand choice, IIA

*This research was supported by a Monash University, Faculty of Business and Economics research grant. We are grateful to the Roy Morgan Research Centre for use of the data. The views expressed in this paper are, however, those of the authors.

1. Introduction.

Determining the factors that influence a household's decision to purchase one brand rather than another is important in understanding the workings of any particular product market. As such a model of brand choice behavior can be a useful tool in brand management. Of interest in modeling such brand choice problems is determining the role played by marketing variables, such as price, and that played by characteristics of the consumer, such as their "loyalty" and demographics. This paper is concerned with estimating a model of brand choice for the laundry detergent market in metropolitan Melbourne. We use data from the Consumer Panel of Australia, collected by the Roy Morgan Research Centre, to estimate and validate our model. This data set provides a rich source of information on both the purchasing behavior and socio-economic and demographic information on the panel members. To our knowledge this is the first such modeling exercise either for this product market or using this data.

The model that we use is a mixed Logit model (see Fry *et al* (1993)). A concern with using a Logit model in such situations is the strong restriction it imposes upon cross elasticities (see Malhotra (1984)). In the Logit model the cross elasticities are constrained to be equal. *A priori* this restriction seems unlikely to be appropriate for marketing applications, such as this modeling exercise. The restriction stems from the Independence of Irrelevant Alternatives (IIA) property. We, therefore, test for IIA using an appropriate test procedure (Small and Hsiao (1984)).

The plan of the rest of this paper is as follows. In section 2 we describe the application and the data set constructed from the Consumer Panel of Australia. Section 3 gives details of the statistical model used and the modeling strategy employed. In section 4 we discuss the results of the analysis and, finally, section 5 contains some concluding remarks.

2. The Application.

The Consumer Panel of Australia (CPA) consists of an Australia wide sample of 2831 households. Data collection is diary based with diaries returned and analyzed monthly. The diaries provide a continuous record of day by day purchases across a large number of product fields. In addition to the purchasing data a wide range of demographic data concerning the household is also collected. The available CPA data concerns the (financial) year July 1992 - June 1993.

In this paper we concentrate upon the laundry detergent market and to avoid potential regional differences consider only the purchases made by the 353 panel members in metropolitan Melbourne. A problem with this market is the existence of a large number of infrequently purchased brands[1]. To reduce the heterogeneity in the data we further restrict our attention to the top 10 purchased brands. However, to ensure that all of the products of the two major manufacturers (Colgate-Palmolive and Unilever) are included in the analysis we define a "brand" termed: Other Unilever. The definition of the brands can be found in Table 1. From this table we see that several of our 11 brands include more than one variant. For example, Cold Power includes both the standard product (Cold Power) and the concentrate version (Cold Power Ultra). The Other Unilever "brand"

[1] The CPA records 99 brands for this product field.

contains 4 Unilever brands and is therefore potentially the most heterogenous outcome.

Thus the data to be analyzed consists of 2283 purchases of the 11 brands by the 353 Melbourne panelists over the period July 1992 - June 1993. This data is then split into two parts: an estimation sample and a holdout (or validation) sample. The estimation sample consists of the 1557 purchases made by the panelists in the first 34 weeks of the sample period and the holdout sample consists of the 726 purchases made by the panelists in weeks 35 - 52.

The outcome to be modeled is the brand j chosen by household i at purchase occasion t. The purchase shares in the estimation sample for the brands comprising this outcome variable are to be found in Table 2. This table shows that these shares vary from 4.04% (Other Unilever) to 13.94% (Omo). However, it is clear that no one brand dominates the market. Additionally neither of the two major manufacturers dominates the market.

Two types of potential explanatory variables are available in the data set. Those that vary across households, i, brands, j, and purchase occasion, t, which will be labeled $z_{ij}(t)$, and those that only vary across household, labeled x_i. Two $z_{ij}(t)$ variables are used here. These concern a measure of price and a measure of brand loyalty constructed from previous purchase history. Potential x_i variables are the household demographics. We now consider the two types of variables in turn.

For any household the only price recorded in the data is that of the purchased brand at purchase occasion t. Estimation of the statistical model requires knowledge of the price of the other 10 brands at purchase occasion t. To estimate the missing prices we use the average price across all purchases of the brand within the same week and from the same store as the actual purchase took place. So, for example, the estimated value of the price of Surf for a household who bought Radiant in a Coles Supermarket in week 7 would be the average price of Surf in a Coles Supermarket in week 7. These prices are all in the form of cents per kilogram. Descriptive statistics for the prices can be found in Table 3. These show that Radiant is the most expensive brand and Spree the cheapest. The "price" brands in this market are Spree, Surf, R.M. Gow and Bushland and the "premium" brands are Dynamo, Drive and Radiant[2]. The price of Other Unilever is the most variable, which is consistent with its definition comprising, as it does, of 4 separate Unilever brands.

The second $z_{ij}(t)$ variable used here concerns brand loyalty. There is much debate in the literature concerning the modeling of loyalty (see *inter alia* Allenby and Lenk (1995)). In this paper we choose to use the loyalty measure introduced by Guadagni and Little (1983). This is defined by:

$$z_{ij}(t) = \lambda z_{ij}(t-1) + (1-\lambda) \begin{cases} 1 \text{ if household } i \text{ bought brand } j \text{ at purchase occasion } \textit{(t-1)} \\ 0 \text{ otherwise.} \end{cases}$$

where λ is a decay parameter. This measure is, therefore, a geometric lag of the households previous purchase history. In this respect it is not unlike the adstock variables often used to model advertising effects (see Broadbent (1979), Broadbent and Fry (1995)). As such we may view the loyalty value for a given brand as representing the households

[2]None of the, so called, "price" brands are advertised. Whereas, the "premium" brands are. The advertised brands in the market are: Cold Power, Dynamo, Drive, Omo, Other Unilever and Radiant.

"commitment" to that brand in the past. The closer to one the value the more loyal the household is to the brand. A value of one would mean that only that brand had been purchased in the past. Conversely, the nearer to zero the less loyal the household and a value of zero would indicate that the household had never bought the brand in the past. A further constraint is that the loyalty measure sums to one across brands. Thus this measure can also be interpreted as representing a share of "commitment" with large values indicating that the brand has featured strongly in the households purchase history.

A choice to be made in using this loyalty measure is the determination of the value of λ. There are two possibilities. It could be calibrated (or estimated) from the data or it could be set, *a priori*, to a given value. In this paper to be consistent with the previous literature, λ is set equal to 0.875 (Guadagni and Little (1983), (1987) and Chintagunta (1993))[3]. The final issue to be addressed in the construction of the loyalty measure is the initialization problem. In other words, what value should $z_{ij}(1)$ take? Our solution is to use the same initialization scheme as Guadagni and Little (1983). That is:

$$z_{ij}(1) = \begin{cases} \lambda \text{ if household } i \text{ bought brand } j \text{ at purchase occasion } \mathit{1} \\ (1-\lambda)/10 \text{ otherwise.} \end{cases}$$

Descriptive statistics for the loyalty measure can be found in Table 3. The mean values of these are similar to those of the purchase shares. It is also clear that within the data set we have purchase occasions where the household is strongly committed to a given brand and quite a large amount of variability in the values taken by this measure.

[3]In terms of an adstock interpretation of this variable $\lambda = 0.875$ corresponds to a "half-life" of 5 purchase occasions.

In what follows we will find it useful to consider two configurations of values for the loyalty measure. The first characterizes a "loyal" household. In this case $z_{ij}(t) = \lambda$ for the brand, j, to which the household is loyal and equals $(1 - \lambda)/10$ for all other brands. The second configuration has all brands featuring equally in the household purchase history. This means that $z_{ij}(t) = 0.0909 = 1/11$, $j = 1, \ldots, 11$. In other words, the household is equally committed to all brands. We call this household a "switcher".

A large amount of demographic information is available in the CPA. This includes information on, *inter alia*, age, sex, education, employment and marital status of household members and pet ownership. In this paper we use variables concerning household size (*HH Size*) and household income (*HH Income*). *HH Size* is the number of persons in the household (minimum 1, maximum 8 and mean 3.5 - see Table 3) and *HH Income* is a grouped version of the total income of the household[4] (minimum 2 ($0-6000), maximum 19 ($100,000+) with a mean of 11 ($35,000-39,999) - see Table 3).

3. The Model.

The dominant modeling paradigm in such brand choice modeling situations is the random utility maximization model (RUM) (see, *inter alia*, Fry *et al* (1993), McFadden (1986)). The model to be estimated in any application depends upon the assumption concerning the random component in the underlying RUM. An assumption of independent Extreme Value distributions leads to a Logit model and an assumption of multivariate Normality

[4]This variable is defined as: 2=$0-6000, 4=$6000-9999, 6=$10000-14999, 7=$15000-19999, 8=$20000-24999, 9=$25000-29999, 10=$30000-34999, 11=$35000-39999, 12=$40000-44999, 13=$45000-49999, 14=$50000-59999, 15=$60000-69999, 16=$70000-79999, 17=$80000-89999, 18=$90000-99999, 19=$100000+.

leads to a Multinomial Probit (MNP) model. Despite recent advances (*e.g.* McFadden (1989), Chintagunta (1992)) the MNP model is computationally burdensome to estimate for a large number of outcomes in the choice set. Given that we have 11 brands (outcomes) in our choice set we therefore choose to use a Logit model in our empirical work. The variant of the Logit model used is termed the mixed Logit model (see Fry *et al* (1993)) and is given below.

In the mixed Logit model the probability that household i will choose brand j on purchase occasion t ($P_{ij}(t)$) is given by:

$$P_{ij}(t) = \frac{\exp\left(\mathbf{z}'_{ij}(t)\boldsymbol{\alpha} + \mathbf{x}'_i\boldsymbol{\beta}_j\right)}{\sum_{k=1}^{11}\left[\exp\left(\mathbf{z}'_{ik}(t)\boldsymbol{\alpha} + \mathbf{x}'_i\boldsymbol{\beta}_k\right)\right]},$$

$i = 1, \ldots, 353$; $j = 1, \ldots, 11$; $t = 1, \ldots, T_i$; $\sum_i T_i = 1557$. The parameters to be estimated are the elements of $\boldsymbol{\alpha}$ and of the $\boldsymbol{\beta}_j$ vectors. In our model specification there are two $z_{ij}(t)$ variables (price and loyalty) and two $\mathbf{x}_i$ variables (household size and household income). In addition to these there are a set of "brand specific" constants. To identify the model $\boldsymbol{\beta}_1 = \mathbf{0}$ and $\boldsymbol{\alpha}$ contains no constants. Thus there are 2 α parameters, 10×2 β parameters and 10 constants (a total of 32 parameters) to be estimated.

To evaluate the model we look for the joint significance of the set of coefficients on *HH Size* and on *HH Income* and of the 10 "brand specific" constants. We also determine the significance of individual coefficients using asymptotic t statistics ($\sim N(0,1)$). *A priori* we expect the coefficient on price to be negative and on loyalty to be positive. We have no *a priori* expectations on the household size and income coefficients nor on the "brand specific" constants. A likelihood ratio test of "overall fit" is also conducted by comparing the maximized log-likelihood of the model[5], $\log L(\hat{\theta})$, with that of a null model, $\log L(\tilde{\theta})$, in which $\hat{P}_{ij}(t)$ is taken to be the purchase share in the estimated sample[6]. To summarize the overall fit we present a pseudo-R^2 measure $(1 - \log L(\hat{\theta}) / \log L(\tilde{\theta}))$. Finally, once the final model has been determined we evaluate its predictive accuracy by examining a "hit-miss" table. This is a tabulation of the predicted versus observed outcomes where the outcome predicted is the value of j for which $\hat{P}_{ij}(t)$ is the maximum.

One, potentially restrictive, property of the Logit model which needs to be tested is that of Independence of Irrelevant Alternatives (IIA). This property states that the ratio of the probability of a household choosing one brand rather than another (*e.g.* Surf rather than Drive) on a given purchase occasion does not depend upon the characteristics, or existence, of any other brands in the choice set. A direct consequence of this is that the cross-elasticities (*e.g.* cross-price) are constrained to be equal (see Hausman (1975), Malhotra (1984)). *A priori* we might expect that this property is unlikely to hold for our application. Since the use of a Logit model depends critically upon this assumption we should therefore test the IIA property. Many tests exist for IIA and in this paper we choose to use the test due to Small and Hsiao (1984), henceforth the SH test, as it has a known asymptotic distribution and has good size properties (see Fry and Harris (1996)).

The SH test procedure involves randomly dividing the estimation sample into two, asymptotically equal, parts A and B, with respective sample sizes n_A and n_B. The Logit model is then estimated over the full choice set, C, in both sub-samples. These estimations

[5]In what follows the parameters of the model are arranged as $\boldsymbol{\theta} = (\boldsymbol{\alpha}', \boldsymbol{\beta}_2', \ldots, \boldsymbol{\beta}_{11}')'$.
[6]Formally the null model is equivalent to a model containing only the 10 "brand specific" constants.

yield two estimates, $\hat{\theta}_C^A$ and $\hat{\theta}_C^B$, which are combined to form $\hat{\theta}_C^{AB}$ $(= \frac{1}{\sqrt{2}}\,\hat{\theta}_C^A + (1-\frac{1}{\sqrt{2}})\hat{\theta}_C^B)$. The sub-sample A is then discarded and the Logit model estimated using sample B for a subset, D, of the full choice set. This estimation yields a maximized log likelihood, $\log L(\hat{\theta}_D^B)$. The SH test statistic is then given by:

$$SH = -2[\log L(\hat{\theta}_C^{AB}) - \log L(\hat{\theta}_D^B)],$$

where $\log L(\hat{\theta}_C^{AB})$ is $\log L$ evaluated at $\hat{\theta}_C^{AB}$. This test statistic has an asymptotic χ^2 distribution with degrees of freedom equal to the dimension of $\hat{\theta}_D^B$ (*i.e.* the identifiable component of θ). In our example we form D by removing R.M. Gow and Bushland from the full set C and thus SH has 26 degrees of freedom.

As mentioned earlier our data set is partitioned into two parts an estimation sample comprising of the 1557 purchases in the first 34 weeks and a holdout sample of 726 purchases in the last 18 weeks. We use the estimation sample to estimate the model parameters, elasticities, to test for the IIA property and to assess the ability of the model to explain the observed choice behavior. Estimation is carried out using maximum likelihood methods with the econometric software package TSP (Hall *et al* (1991)). The holdout sample is used to assess the predictive accuracy of the estimated model and, therefore, provides an additional check on the model validity. The results of the modeling are discussed in the next section.

4. Results.

Table 4 contains the results of the maximum likelihood estimation of the Logit model. We see that the signs of both the price and loyalty variables are in line with our *a priori* expectations and are highly significant. Joint tests for the size variable, income variable and brand constants confirm them to be significant explanatory factors. The overall LR test provides a strong rejection of the null model confirming that the fitted model is adequate at explaining the observed choice behavior. This is supported by the, fairly high, pseudo-R^2 value of 0.545. The SH test procedure for the IIA property yields a test statistic value of 28.6 which is below the 5% critical value of the χ^2 distribution with 26 degrees of freedom. Thus the IIA property appears to be supported in this application. This finding concerning IIA is, perhaps, surprising, but may reflect the composition of the choice set.

We now turn to an interpretation of the estimated model. It is clear that loyalty is an extremely important determinant of brand choice. Unfortunately, given its construction, it is difficult to directly interpret responses to changes in the loyalty measure (*e.g.* loyalty elasticities). Instead we concentrate on the impact of loyalty on the estimated price elasticity. That is we present two sets of price elasticities: one for "loyal" households and one for "switchers". These estimated elasticities are found in Table 5. We see that the own price elasticity for a "switcher" is five, or more, times that of a "loyal" household. "Loyal" households are price inelastic for all brands. "Switchers" are price inelastic for the, so called, "price" brands and price elastic for the, so called, "premium" brands. The cross price elasticities are all equal, as constrained by the IIA property, and are higher for "loyal" households. The cross elasticities also show a clear distinction between the "price" and "premium" brands. These findings are in line with expectations.

Increasing household size, *ceteris paribus*, increases the probability[7] of a household choosing Fab, Spree, Drive, Other Unilever, R.M. Gow and Bushland and decreases the probability of choosing Dynamo, Omo, Surf, Radiant. Similarly increasing household income, *ceteris paribus*, increases the probability of choosing Dynamo, Spree, Omo, R.M. Gow and Radiant and decreases the probability of choosing Fab, Drive, Surf, Other Unilever and Bushland. There appears to be no pattern in these results which suggests any particular structure in the choice set.

The "brand specific" constants would be zero if the model included all of the explanatory factors needed to explain the observed choice behavior. The fact that they are not zero suggests that unique factors specific to the brands exist, which these constants are modeling. In common with Guadagni and Little (1983) we rank these coefficients from smallest to largest and look to see whether they are related to any known factors (see Table 6). Unlike their findings, we do not see a relationship between the rank of the constants and the ranked purchase shares. We do, however, see a potential relationship between the constants and the "price" and "premium" brands and between the constants and the use of advertising. The constants for the "price" brands are predominately negative and are all positive for the "premium" brands. The relationship with the presence of advertising is also reflected in positive constants. These results suggest that any potential excluded factor(s) in the model would relate to the distinction between "price" and "premium"

[7]The increase or decrease in selection probability is relative to the normalization. In this case relative to the probability of choosing Cold Power.

brands[8].

Using the estimated model we form $\hat{P}_{ij}(t)$, for each i, j, t, and use these to predict the brand that would be chosen. The predicted brand is the one for which $\hat{P}_{ij}(t)$ is the maximum. The predicted and observed choices are then cross tabulated in a "hit-miss" table. If the model is a perfect predictive model then it would always predict the observed choice and all observations in the table would lie on the main diagonal. It should be noted that this procedure is a separate validation of the model as the objective function used in estimation (the log-likelihood) is not necessarily the one which would optimize predictive accuracy. The results of this exercise are found in Table 7. Overall 64.6% of observed choices were correctly identified by the model. The predictive accuracy varied from 50.8% for Other Unilever to 74.8% for Dynamo. There was also a tendency for incorrect predictions for "price" brands to be other "price" brands (*e.g.* 12% of Surf purchases were predicted to be purchases of Spree). Such a pattern was not apparent for the "premium" brands.

The final check on the estimated model involves the use of the holdout (or validation) sample. We use our estimated model to predict the brand choices for the holdout sample. That is we take the data on the $z_{ij}(t)$ and x_i variables and using $\hat{\theta}$ find $\hat{P}_{ij}(t)$. This procedure requires knowledge of the values of the loyalty measure. For the first purchase of any household we initialized the loyalty measure as though the household was a "switcher" ($z_{ij}(1) = 0.0909$, $j = 1, \ldots, 11$) and thereafter used the observed purchase history to update the measure. The results of this validation exercise are found in Table 8. Overall

[8] Unfortunately, data on advertising expenditures is not available.

75.2% of observed choices are correctly predicted. The variation was between 60.8% correct predictions for Drive to 90.5% correct predictions for Radiant. Once again the pattern of incorrect predictions for a "price" brand being other "price" brands is present.

5. Conclusions.

In this paper we have fitted a mixed Logit model of brand choice to data on purchases of laundry detergent in metropolitan Melbourne as recorded by the Consumer Panel of Australia in the year 1992/93. We find that price, brand loyalty, household size, household income and a set of "brand specific" constants are all important determinants of brand choice behavior. Price elasticity is affected by brand loyalty with price elasticities being much smaller for brand loyal households. There appears to be a segmentation of the market into "price" and "premium" brands suggested by our results.

Overall the model fits well and is good at explaining observed behavior. The IIA property inherent in the use of Logit models appears to be supported by the data. This suggests that the structure of the choice set in this application is consistent with the IIA property. This result is contrary to, *a priori*, beliefs about such brand choice problems. The test result could, however, reflect the fact that the SH test procedure used to test IIA does not have good power properties in this particular application. Additionally, we find that our Logit model performs well in predicting choice behavior in a holdout sample.

Our results suggest that in this application the mixed Logit model does a good job at both explaining and predicting brand choice behavior. This is encouraging given that this is the first modeling exercise carried out using this data set and the results have given some insights into brand choice in this market. Further work might, however, investigate four issues in more detail. These are the subject of brand loyalty effects (Allenby and Lenk (1995)), parameter and individual heterogeneity (Chintagunta *et al* (1991)), the use of non-IIA models such as the Multinomial Probit (Chintagunta (1992)) and Bayesian Probit models (McCulloch and Rossi (1994)), and the exploitation of the true panel nature of the CPA data.

References.

Allenby, G.M. and P.J. Lenk, (1995), "Reassessing Brand Loyalty, Price Sensitivity, and Merchandising Effects on Consumer Brand Choice", *Journal of Business & Economic Statistics*, **13**, 281-289.

Broadbent, S., (1979), "One Way TV Advertisments Work", *Journal of the Market Research Society*, **21**, 139-166.

Broadbent, S. and T.R.L. Fry, (1985), "Adstock Modelling for the Long Term", *Journal of the Market Research Society*, **37**, 385-403.

Chintagunta, P.K., (1992), "Estimating a Multinomial Probit Model of Brand Choice using the Method of Simulated Moments", *Marketing Science*, **11**, 386-407.

Chintagunta, P.K. Jain, D.C. and N.J. Vilcassim, (1991), "Investigating Heterogeneity in Brand Preferences in a Logit Model for Panel Data", *Journal of Marketing Research*, **28**, 417-428.

Fry, T.R.L., Brooks, R.D., Comley, B.R. and J. Zhang, (1993), "Economic Motivations for Limited Dependent and Qualitative Variable Models", *Economic Record*, **69**, 193-205.

Fry, T.R.L. and M.N. Harris, (1996), "A Monte Carlo Study of Tests for the Independence of Irrelevant Alternatives Property", *Transportation Research, Series B*, **30**, 19-30.

Guadagni, P.M. and J.D.C. Little, (1983), "A Logit Model of Brand Choice Calibrated on Scanner Data", *Marketing Science*, **2**, 203-238.

Guadagni, P.M. and J.D.C. Little, (1987), "When and What to Buy: A Nested Logit Model of Coffee Purchase", Working Paper 1919-87, Sloan School of Management, MIT.

Hall, B.H., Schnake, R. and C. Cummins, (1991), *Time Series Processor: Reference Manual, Version 4.2*, TSP International, Palo Alto.

Hausman, J.A., (1975), "Project Independence Report: An Appraisal of U.S. Energy Needs up to 1985", *The Bell Journal of Economics*, **6**, 517-551.

Malhotra, N.K., (1984), "The Use of Linear Logit Models in Marketing Research", *Journal of Marketing Research*, **21**, 20-31.

McCulloch, R. and P.E. Rossi, (1994), "An Exact Likelihood Analysis of the Multinomial Probit Model", *Journal of Econometrics*, **64**, 207-240.

McFadden, D., (1986), "The Choice Theory Approach to Market Research", *Marketing Science*, **5**, 275-297.

Small K.A. and C. Hsiao, (1984), "Multinomial Logit Specification Tests", *International Economic Review*, **16**, 471-486.

Table 1: **Definition of brands.**

Brand			Manufacturer
Cold Power	1	Cold Power, Cold Power Ultra	Colgate Palmolive
Dynamo	2	Dynamo, Dynamo Concentrate	Colgate Palmolive
Fab	3	Fab	Colgate Palmolive
Spree	4	Spree	Colgate Palmolive
Drive	5	Drive, Drive Power	Unilever
Omo	6	Omo, Omomatic, Omo Micro, Omo Free	Unilever
Surf	7	Surf	Unilever
Other Unilever	8	Aura, Down to Earth, Lux, Rinso	Unilever
R.M. Gow	9	Gow's, Ease	R.M. Gow
Bushland	10	Bushland	Bushland
Radiant	11	Radiant	Cussons

Table 2: **Purchase Shares - Estimation Sample.**

Brand	% Share
Cold Power	8.92
Dynamo	9.44
Fab	7.19
Spree	8.09
Drive	12.40
Omo	13.94
Surf	13.36
Other Unilever	4.04
R.M. Gow	7.78
Bushland	5.46
Radiant	9.38

Total number of purchases: 1557.

Table 3: **Descriptive Statistics - Estimation Sample.**

Price	Mean	Std. dev.	Min.	Max.
Cold Power	382.034	79.016	166.333	595.000
Dynamo	418.809	71.084	239.200	838.500
Fab	349.437	67.601	256.200	499.000
Spree	231.644	42.756	132.667	449.000
Drive	411.936	78.614	239.200	639.000
Omo	386.839	83.417	167.000	598.667
Surf	252.857	34.083	145.000	400.000
Other Unilever	431.561	118.867	73.000	564.000
R.M. Gow	232.944	34.864	169.000	329.000
Bushland	172.997	23.574	128.333	267.500
Radiant	446.999	74.075	298.000	718.000

Loyalty	Mean	Std. dev.	Min.	Max.
Cold Power	0.093	0.227	0.000228	0.975
Dynamo	0.094	0.237	0.000228	0.981
Fab	0.068	0.191	0.000068	0.957
Spree	0.093	0.216	0.000068	0.944
Drive	0.118	0.251	0.000340	0.975
Omo	0.140	0.271	0.000340	0.981
Surf	0.130	0.257	0.000068	0.967
Other Unilever	0.038	0.133	0.000068	0.936
R.M. Gow	0.078	0.221	0.000068	0.997
Bushland	0.064	0.189	0.000068	0.957
Radiant	0.084	0.223	0.000068	0.962

	Mean	Std. dev.	Min.	Max.
HH Size	3.541	1.506	1	8
HH Income	11.159	3.523	2	19

Table 4: Estimation Results.

	Variable		
		Price	Loyalty
		-0.00352	4.85449
		(-6.2042)	*(48.3081)*
Brand	**Constant**	**HH Size**	**HH Income**
Cold Power	*	*	*
Dynamo	0.53956	-0.38692	0.08424
	(0.8307)	*(-2.5460)*	*(1.4746)*
Fab	-0.39526	0.13135	-0.00818
	(-0.6468)	*(1.0673)*	*(-0.1421)*
Spree	-2.04765	0.06600	0.10408
	(-3.4170)	*(0.5505)*	*(1.9347)*
Drive	0.44410	0.02943	-0.00809
	(0.7689)	*(0.2499)*	*(-0.1608)*
Omo	-0.03833	-0.27303	0.11849
	(-0.0677)	*(-2.2685)*	*(2.3554)*
Surf	0.50838	-0.17675	-0.00490
	(0.8966)	*(-1.4671)*	*(-0.0949)*
Other Unilever	-0.00614	0.05783	-0.01198
	(-0.0088)	*(0.4147)*	*(-0.1859)*
R.M. Gow	-1.26088	0.04421	0.04928
	(-1.8954)	*(0.3375)*	*(0.8064)*
Bushland	-1.43923	0.16076	-0.03742
	(-1.8420)	*(1.2674)*	*(-0.5906)*
Radiant	0.65002	-0.05213	0.00476
	(1.0614)	*(-0.3970)*	*(0.0880)*

N.B. asymptotic t values in parentheses.

log-L	-1661.29	
Restricted log-L	-3647.54	
LR test	3972.5	$\chi^2(22)$
Pseudo-R^2	0.545	
Small-Hsiao IIA test	28.6	$\chi^2(26)$
Predictive Accuracy	64.6%	

Table 5: **Estimated Price Elasticities.**

Brand	**Own Price** *Loyal*	*Switcher*	**Cross Price** *Loyal*	*Switcher*
Cold Power	-0.200	-1.166	1.213	0.107
Dynamo	-0.224	-1.273	1.331	0.115
Fab	-0.169	-1.063	1.122	0.109
Spree	-0.133	-0.730	0.706	0.059
Drive	-0.154	-1.209	1.384	0.165
Omo	-0.152	-1.146	1.287	0.147
Surf	-0.099	-0.764	0.824	0.096
Other Unilever	-0.251	-1.318	1.351	0.108
R.M. Gow	-0.123	-0.728	0.722	0.065
Bushland	-0.142	-0.564	0.478	0.029
Radiant	-0.179	-1.315	1.497	0.167

Table 6: **Brand Specific Constants and Other Factors.**

Brand	**Rank**	**Share**	**"Price"**	**"Premium"**	**Advertised**
Cold Power	7	6			×
Dynamo	10	8		×	×
Fab	4	3			
Spree	1	5	×		
Drive	8	9		×	×
Omo	5	11			×
Surf	9	10	×		
Other Unilever	6	1			×
R.M. Gow	3	4	×		
Bushland	2	2	×		
Radiant	11	7		×	×

Table 7: **Predictive Accuracy - Estimation Sample.**

Observed		1	2	3	4	5	6	7	8	9	10	11	*Total*
		Predicted											
Cold Power	1	90	6	2	5	9	12	6	4	2	1	2	139
Dynamo	2	1	110	5	7	10	6	1	0	0	1	6	147
Fab	3	6	2	70	6	3	8	6	1	2	5	3	112
Spree	4	4	0	8	71	1	6	18	0	8	8	2	126
Drive	5	7	14	8	7	121	21	3	2	3	5	2	193
Omo	6	16	7	2	9	16	140	15	2	2	2	6	217
Surf	7	7	2	8	25	12	11	125	0	10	3	5	208
Other Unilever	8	4	1	1	0	2	12	8	32	1	2	0	63
R.M. Gow	9	2	1	0	9	2	3	11	0	88	5	0	121
Bushland	10	1	2	0	9	3	2	7	0	0	61	0	85
Radiant	11	5	0	5	4	8	5	11	6	0	4	98	146
Total		143	145	109	152	187	226	211	47	116	97	124	1557

Table 8: **Predictive Accuracy - Holdout Sample.**

Observed		1	2	3	4	5	6	7	8	9	10	11	*Total*
		Predicted											
Cold Power	1	62	2	0	1	4	3	2	0	0	0	1	75
Dynamo	2	2	52	2	2	2	1	1	0	0	0	0	62
Fab	3	0	0	34	1	4	1	9	0	0	0	1	50
Spree	4	3	1	1	43	0	0	20	0	1	0	3	72
Drive	5	4	10	7	1	57	7	4	0	1	0	3	94
Omo	6	0	6	2	4	5	68	9	1	0	0	1	96
Surf	7	1	7	4	5	2	2	77	0	2	0	5	105
Other Unilever	8	0	0	0	0	0	0	2	23	0	1	0	26
R.M. Gow	9	0	0	1	2	0	0	1	0	31	0	1	36
Bushland	10	0	0	0	1	0	0	0	0	1	42	3	47
Radiant	11	0	1	0	4	1	0	0	0	0	0	57	63
Total		72	79	51	64	75	82	125	24	36	43	75	726