Introduction

This assignment will be evaluating and analyzing choice data from qualitative research conducted by Star Technologies Company. The objective will be to provide the product development team with more information on which attributes impact tablet preferences. In addition to qualitative research, STC also collected quantitative research that included choice-based-conjoint (CBC) task. Neverending Marketing insights, a global marketing research services provider, was also hired to obtain a sample of tablet owners, and likely buyers, to program and host an online questionnaire.

Data

We were provided with consumer survey data from 424 respondents. The survey was designed and administered by STC and by Neverending Marketing Insights. The survey included questions focused on consumer interests, choice sets, and demographics. The set of survey questions were:

- Previous owner of an STC Product?
- 36 Set questions
- Interest Questions:
  - Purchasing a new tablet?
  - Purchasing a new smart phone?
  - Using cloud storage for storing personal digital content?
  - Taking an online course to improve relevant skills?
- Gender of respondent
- Age of respondent
- Are you a parent?
- What are the ages of your children?
- Household income
- State of residence

The choice set questions were built form the following five attributes:

- **Brand** - 4 levels: STC, Somesong, Pear, Gaggle (level codes: 0,1,2,3)
- **Price** - 3 levels: $199, $299, $399 (levels: 0,1,2)
- **Screen** - 3 levels: 5-inch, 7-inch, 10-inch (levels: 0,1,2)
- **RAM** - 3 levels: 8 Gb, 16 Gb, 32 Gb (Gb = “gigabytes”) (levels: 0,1,2)
- **Processor**- 3 levels: 1.5 GHz, 2 GHz, 2.5 GHz (GHz = “gigahertz”) (levels: 0,1,2)

Each of the choice set questions had three alternatives with each representing a specific combination of the above attribute levels. Each consumer was to rank the three alternatives presented for each choice set with a score from one to three. We were also given two additional datasets. The first is a dataframe of the 108 different combinations for the choice set questions. The second data frame contains 12 alternatives based on the choice sets of four additional respondents. The first dataframe is used to form the predictor variable set for model estimation. The second is for subsequent preference share modelling.

Data Preparation

We conducted a number of pre-process procedures to manipulate the data into a format which can easily allow for the fitting of an MNL regression model.

1. Create a predictor variable.
2. Create an effects matrix.
3. Create a brand matrix to coincide with brand from the effects matrix.
4. Create a price vector of the difference between the price associated with each attribute combination and the mean price of all possible attribute combinations.
5. Create an X-matrix to merge the effects matrix and the brands by price matrix.
6. Create a list of y by X-matrix which is with be a combination of each respondents' selection to the 36 choice questions (y-matrix) and the matrix of predictor variables (x-matrix).

Model 1

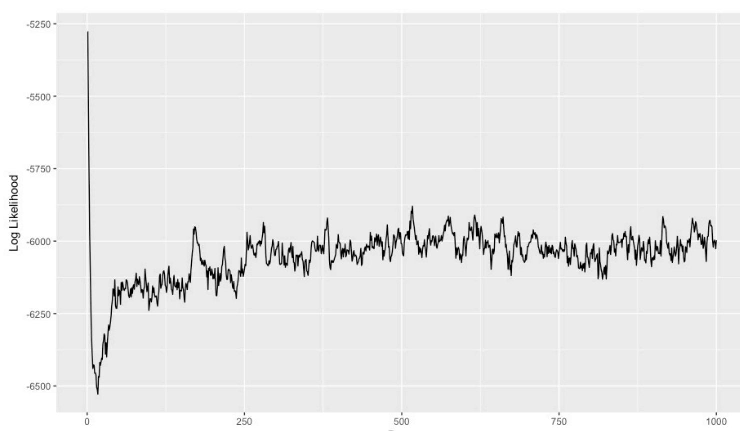
The first model that was created included the choice set responses for all 424 respondents along with the predictor variables in the (X-matrix). This MNL likelihood model can be shown as the following formulas:

$y_{ijk} = X_{ijk} * \beta_i + e_{ijk}$

Where:

$\beta_i \sim MVN(\bar{\beta}, V_\beta)$

This model was used to ensure that we have included enough iterations for proper use of coefficients and then that we assess the beta means directly to determine whether or not each seem logical. In order to calculate statistical measures that would allow us to determine fit we needed to find the posterior means for each beta of each respondent. After, a matrix would be created from multiplying the original X-matrix by the posterior means. Then, it would then be exponentiated and divided by the row of sums of the exponentiated betas. The log likelihood values recorded over the MCMC simulation for Model 1 are shown below. You will note that the beta and likelihood value tend to level out around 500 iterations. This would suggest that we included more than enough iterations to achieve a proper "burn-in".



The mean betas can be found below. We estimated the beta mean for the reference level by solving the sum of beta mans for all levels of that attribute for zero. Solving for zero is appropriate because our specification does not have an intercept term. From the data that we currently have in the beta means, we can note the respondents have a preference for a 10” screen, 32 Gb of RAM, 2.5 GHz processor speeds, and a $199 price point, and that the tablet is of the STC brand. The attribute preference for screen size, RAM and processor speed are reasonable, noting that greater levels for these attributes coincide with values which provide more functional value. The preference for a lower price is an obvious preference for a consumer. As the results currently stand, we have little justification for the preference of the STC brand. However, we can note that the small magnitude of coefficient for the Somesong, Pear, and Gaggle brand attribute.

Attribute	Beta Mean
7” Screen	-0.1415
10” Screen	0.4327
16Gb RAM	0.0611
32Gb RAM	0.6202
2 GHz Processor	0.9688
2.5 GHz Processor	1.2489
$299 Price	0.2669
$399 Price	-2.7071
Somesong	-0.1153
Pear	0.0367
Gaggle	-0.2306
Somesong Brand by Price	0.1496
Pear Brand by Price	0.0083
Gaggle Brand by Price	-0.0132

Model 2

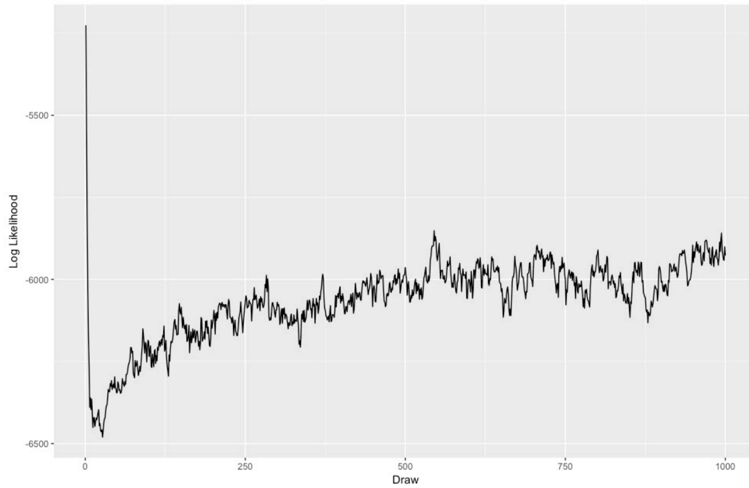
The second model that we ran was fitting an HB MNL model with prior ownership as a covariate. This model also includes the choice set responses for all 424 respondents along with the predictor variables. Additional covariates of responses to the survey question STC Owner was included. We show the beta’s as dependent variables of the covariates as:

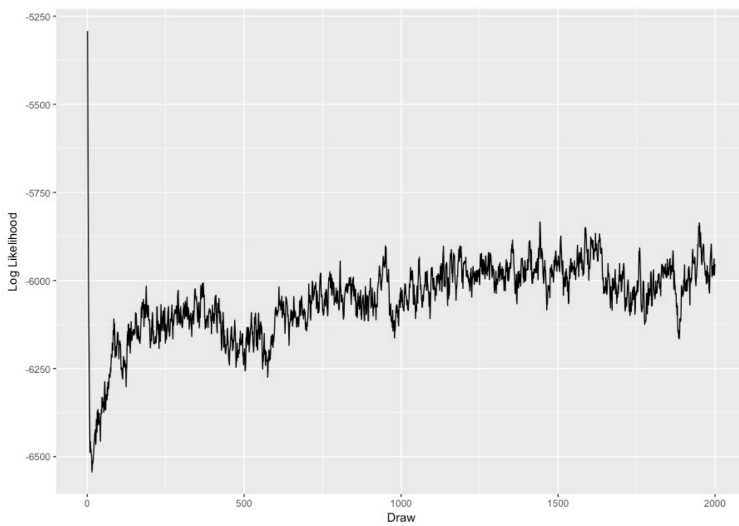
$\beta_i = \psi * Z_i + u_i$

Where:

$u_i \sim MVN(0, V_u)$

The first plot of the Log Likelihood values recorded over the MCMC simulation for model 2 can be seen below (First). From the looks of the plot, it does not look like we did enough iterations since at the 1,000 level the Log Likelihood plot does not level off. To fix this issue, we re-ran the iterations increasing the number of iterations from 5,000 to 10,000.





The second Loglikelihood plot shows the plot leveling off around 1500 and indicates that we have performed enough iterations.

Attribute	Beta Mean	STC Owner Delta Mean	Gender Delta Mean
7” Screen	-0.1415	-0.0381	-0.1887
10” Screen	0.4327	-0.0123	0.4682
16Gb RAM	0.0611	0.0971	0.0780
32Gb RAM	0.6202	0.0340	0.6072
2 GHz Processor	0.9688	0.2244	0.9907
2.5 GHz Processor	1.2489	0.4345	1.2654
$299 Price	0.2669	-0.0512	0.2720
$399 Price	-2.7071	-0.5334	-2.7277
Somesong	-0.1153	-0.1978	-0.1596
Pear	0.0367	0.8967	0.0993
Gaggle	-0.2306	0.0368	-0.2555
Somesong Brand by Price	0.1496	-0.2152	0.0650
Pear Brand by Price	0.0083	0.0489	0.0915
Gaggle Brand by Price	-0.0132	0.2529	0.0037

After the code is ran we can assess the delta means of the STC Owner covariate in order to determine whether or not prior ownership of an STC product has an impact on attribute preferences. So far, we can surmise that those who have previously purchased an STC branded product have a greater preference for a 5” screen, 8 Gb of ram, a 2.5 GHz processor, and a $199 price point. It is also interesting to note that those who had previously purchased an STC product had a greater preference for Pear branded products. To see what we got for results, we also included Gender to see if that would make a difference for performance attributes. We noticed that there is a similar bias for males to have greater preference toward tablets with a 10-inch screen, 32 Gb of RAM, a 2.5GHz Processor, the $199 price point, and the STC brand.

Customer Choice Scenarios

We can use estimated models as well as the customer choice prediction model to estimate preference shares for the alternatives in each of the additional scenarios. We created an X-matrix which includes an effects coded version of the additional provided scenarios. We calculated the posterior means for each beta for each subject by using the betas from model 1. Then, we multiplied this matrix by the transpose of the mean betas matrix. The choice probabilities in order will allow us to derive a set of preference shares for each attribute. We calculated the conjoint part-worth utilities for each respondent which is found by dividing the sum of the number of times an attribute was selected by the number of times the attribute was available for selection. The ratio gives an indication of attribute preference and will be used to measure the relative importance of those attributes.

Attribute	Share
7” Screen	-0.1046
10” Screen	0.3355
16Gb RAM	0.1380
32Gb RAM	0.5466
2 GHz Processor	0.9653
2.5 GHz Processor	1.1701
$299 Price	0.2459
$399 Price	-2.7078
Somesong	-0.1706
Pear	0.1447
Gaggle	-0.2369
Somesong Brand by Price	0.1184
Pear Brand by Price	0.1086
Gaggle Brand by Price	-0.0112

From looking at the preferences of individual users, we can conclude that there is a higher preference for a tablet with a 10” screen, 32 Gb of ram, a 2.5 GHz Processor, the $299.00 price point of the STC brand.

Conclusion

This assignment was conducted to fit Hierarchical Bayes Multinominal Logit models in order to measure the sensitivity of respondent choices relating to attributes of the products. We also examined whether or not prior STC product ownership affected attribute and product preference. An evaluation of the means betas for brand by price interaction terms of Model 1 showed us that there was little variation between mean beta values and little price sensitivity between different brands. This would indicate that it was positive that STC entered the market. After reviewing Model 1, we noticed that respondents had a preference for tablets with a 10 inch screen, 20 Gb of RAM, a 2.5GHz processor, and a $199 price point.

In Model 2, we evaluated whether or not prior STC product ownership affected brand preference. We found that those who had previously owned an STC product tended to have a strong preference for Pear branded products. STC would do well to target Pear branded products in their advertising campaigns when they enter the market. Offering a cheaper alternative to Pear branded products may bear some fruit when it comes to 2.5 GHz Processor models with 32Gb of RAM.

As far as attribute preference is concerned, STC would benefit from providing two different product lines. The first product line would include a larger screen size, greater amount of RAM, and a faster processor as a higher, yet still competitive, price point. The second would be a bargain option with a smaller screen size, less RAM, and a slower processor at a much lower price point. The first, more expensive product, should be targeted towards Pear brand loyal customers. The marketing campaign to introduce both of these products should be targeted against Pear branded products.