Choice-Based Conjoint Analysis
Star Technologies Company

Northwestern University
Andrew Kang

Introduction

Star Technologies Company is conducting a Choice-Based Conjoint Analysis in order to determine what product they should go to market with. In the survey conducted, each respondent was given 36 designs and 3 alternatives.

The purpose of our model is to interpret these results and assess the impact that product attributes have on the consumer's preferences. With our model, we seek to understand attribute importance and how best use our research to drive product recommendations.

Data Preparation

In preparing our model, we use effect coding to capture the 108 combinations that can be derived from the 36 choice set questions within a matrix. We also add to additional columns to our matrix for each brand and include the price level associated with each brand for the 108 combinations. We then capture the responses from the survey participants for each choice set presented into a separate matrix.

Modeling

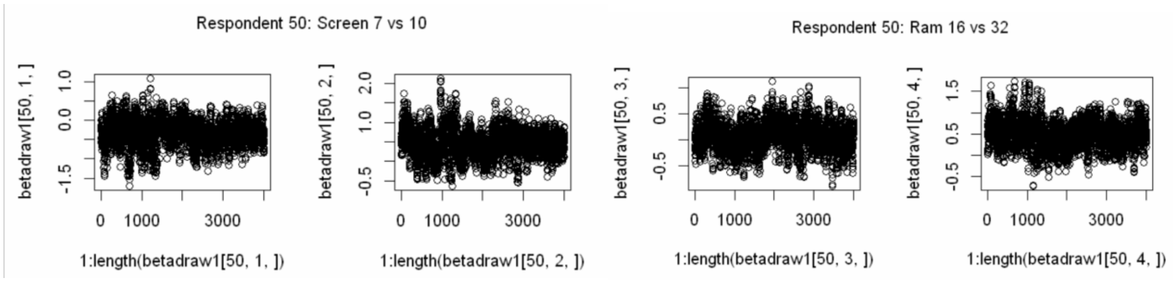
The Hierarchical Bayes Multinomial Logit regression model uses a Bayesian approach of assuming a prior distribution and multiplying it against the likelihood to achieve a new posterior distribution. Monte Carlo Markov Chain simulation allows for estimation regarding the shape of the distribution of the posterior by generating random points and evaluating whether those points landed in either the rejection region or acceptance region. The method we employ, Metropolis-Hastings, obtains random samples from the probability distribution and takes steps towards higher likelihood while randomly staying or moving when likelihood is less. The concept here is that over a large enough number of iterations, the simulation can converge closely to the actual results over the long run.

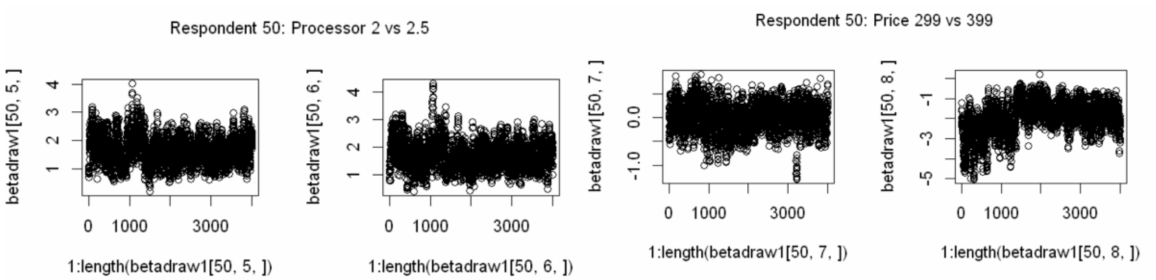
Model 1: Model Interpretation

For our MCMC model, we used 20,000 iterations and took one out of every 5 draws. We attempted to perform more iterations, but the models would fail at higher iterations due to warnings related to matrix symmetry warnings that were being output to the R console.

When analyzing our results, we see that we have some inconsistency related to convergence. Through sampling a few of our respondents, we see that we have some convergence occurring after 2000 data points. We want to preserve as much data while also ensuring that we allow for a proper burn-in, so we will throw away the first 2,000 points. This equates to a 10,000 iteration burn-in. Plots for respondent 50 can be seen below.

Burn-In Charts



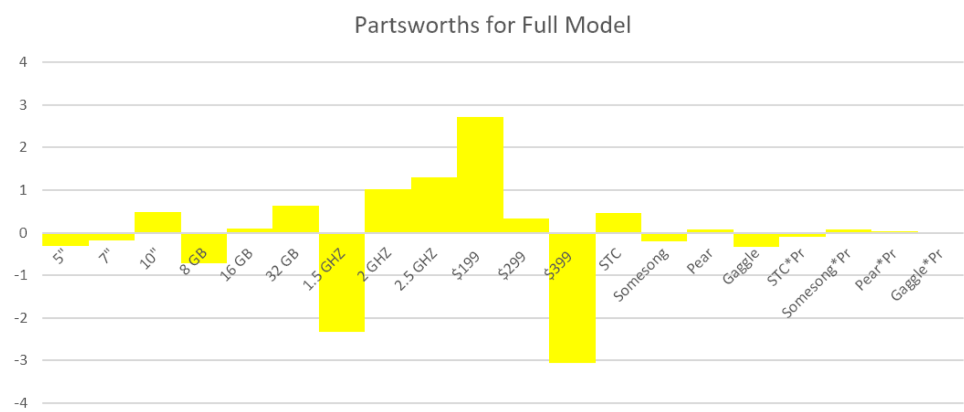


Following our burn-in selection for iterations, we can now evaluate the samples from our posterior distributions, capture betameans coefficients, and calculate the odds ratio. With these calculations, we can compare different levels of attributes to extract preference.

Betameans and Odds Ratio Calculations

Variable	Level	Coef	Log(OR)	Odds Ratio
Screen	5"	Base	-0.316628773	0.728601181
Screen	7"	x1	-0.173545366	0.840679012
Screen	10"	x2	0.49017414	1.632600495
RAM	8 GB	Base	-0.716690268	0.488365945
RAM	16 GB	x3	0.090188221	1.094380249
RAM	32 GB	x4	0.626502047	1.871054259
Processor	1.5 GHZ	Base	-2.31831599	0.098439219
Processor	2 GHZ	x5	1.017943014	2.767496205
Processor	2.5 GHZ	x6	1.300372976	3.670665481
Price	$199	Base	2.712051372	15.06013779
Price	$299	x7	0.342995402	1.409162283
Price	$399	x8	-3.055046774	0.047120517
Brand	STC	Base	0.468441014	1.597501768
Brand	Somesong	x9	-0.211828685	0.809103297
Brand	Pear	x10	0.077590075	1.08067957
Brand	Gaggle	x11	-0.334202404	0.715908865
Brand*Price	STC*Pr	Base	-0.102574362	0.902511035
Brand*Price	Somesong*Pr	x12	0.078455467	1.081615186
Brand*Price	Pear*Pr	x13	0.043866155	1.044842499

Brand*Price	Gaggle*Pr	x14	-0.01974726	0.98044644



In terms of screen size, we see that people prefer a low price point at $199 15 times more than not preferring. In terms of RAM, people prefer 1.87 times more than not preferring. In terms of screens, people prefer 10-inch screens 1.6 times more than not preferring. In terms of brand preference, people prefer STC 1.6 times more than no preference.

Overall, consumers prefer a $199 price, 2.5 GHZ processor, 10-inch screen, and the STC brand. What remains to be seen is whether price and brand have interaction effects and whether we can further segment our respondents based on ownership of STC products.

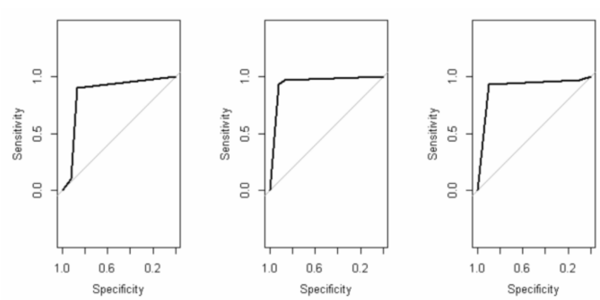
Confusion Matrix

In the confusion matrix below, we see that our model performs quite well at predicting consumers that will make choices 1,2, and 3. Our accuracy is 85.35%.

```
          ydatavec
custchoice    1    2    3
         1 3634  468  212
         2  249 3673  229
         3  321  448 6030
```

Multiclass Specificity vs. Sensitivity Plots

In terms of AUC, our model has a metric of 90.04%, which denotes that our model's classifer performance averaged across different classifier parameters.



Price Sensitivity for Brands

Given our computations for our full model, we can now interpret brand preference relative to price. In order to do so, we take the log odds ratio and raise it to the exponent to extract the odds ratio.

The table below shows the odds of preference for different price points. We see that our product is strong in that we can command strong preference relative to other brands at all price points. What we notice, is that as price increases, our odds ratio goes down except against Pear. For Pear, the odds ratio increases, which suggests that we may want to evaluate a strategy against Pear depending on their market share. We have preference throughout the prices listed, so we should try and maximize market share depending on who is the market leader.

		$199	$299	$399
Brand 1	**Brand 2**	**-1**	**0**	**1**
STC	Somesong	2.366233682	1.97441	1.647469
STC	Pear	1.3030684	1.338075	1.374022
STC	Gaggle	2.18303755	2.178249	2.173472

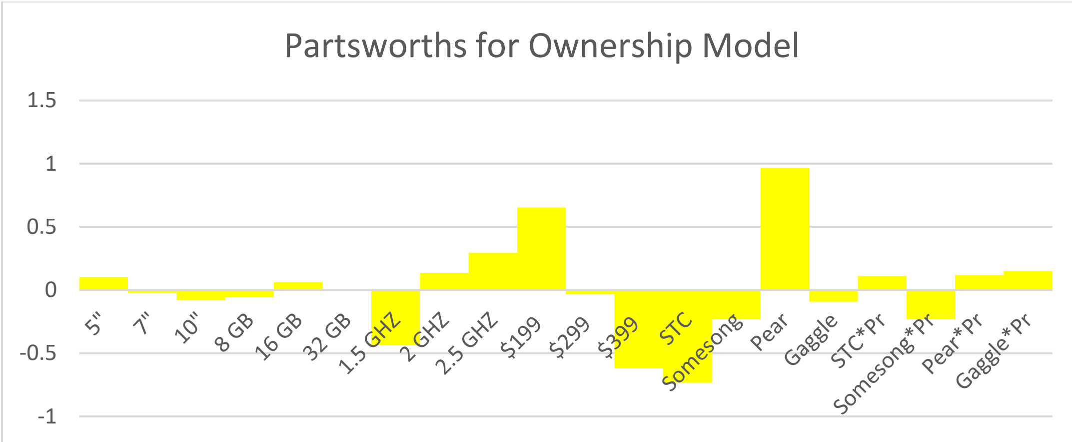
Model 2: Model Interpretation

Impact of Ownership on Preferences

In our second model, we seek to answer the question regarding whether our results are correlated with ownership. In order to evaluate this, we take a new metric, deltadraw, and extract out betadraw again. By evaluating the directional deviations from our initial model, we will be able to identify potential risk areas from existing customers that may affect customer lifetime value.

Betameans, Odds Ratio, and Deltameans Calculations

Variable	Level	Coef	Log(OR)	Odds Ratio	Deltameans
Screen	5"	Base	-0.304032989	0.737836526	0.102118455
Screen	7"	x1	-0.188467271	0.82822761	-0.02071216
Screen	10"	x2	0.49250026	1.636402541	-0.08140629
RAM	8 GB	Base	-0.72718082	0.483269495	-0.05772741
RAM	16 GB	x3	0.123428929	1.131369594	0.062044541
RAM	32 GB	x4	0.603751891	1.828968032	-0.00431713
Processor	1.5 GHZ	Base	-2.294103787	0.100851737	-0.4332764
Processor	2 GHZ	x5	1.018304921	2.76849796	0.137096432
Processor	2.5 GHZ	x6	1.275798866	3.581561456	0.296179965
Price	$199	Base	2.638025938	13.98556796	0.65246176
Price	$299	x7	0.329663385	1.390499987	-0.03329167
Price	$399	x8	-2.967689323	0.051421993	-0.61917009
Brand	STC	Base	0.446406636	1.562686783	-0.73393623
Brand	Somesong	x9	-0.185220023	0.830921442	-0.23188307
Brand	Pear	x10	0.074829877	1.077700794	0.965819305
Brand	Gaggle	x11	-0.336016491	0.714611321	-0.09163566
Brand*Price	STC*Pr	Base	-0.118571064	0.888188695	0.109656538
Brand*Price	Somesong*Pr	x12	0.097970365	1.102930099	-0.22885065
Brand*Price	Pear*Pr	x13	0.022476902	1.022731411	0.119194109
Brand*Price	Gaggle*Pr	x14	-0.001876202	0.998125557	0.151022477



In our second model, we see that deltameans reflects the positive or negative sentiment surrounding purchases after owning an STC. What we notice form our analysis is that consumers have a greater affinity for 5" screens, 2.5 GHZ processors, and less likely to repurchase STC and more in favor of Pear.

Given the likelihood that STC consumers will switch to Pear, it seems very likely that consumers in our segment see Pear as a comparable alternative to STC. In our initial model, we saw that 10-inch screens were preferred whereas the STC consumers want 5-inch screens

Given the strength of the Pear brand, $199 prices and 2.5 ghz processor, we could see that in fact the ownership covariate does influence attributes and preferences.

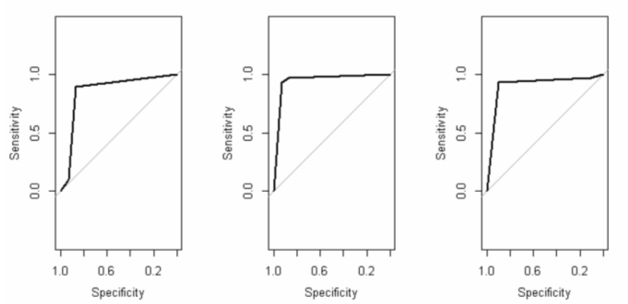
Confusion Matrix

Our confusion matrix in model shows similar performance to our first model. We see, however, that we do increase our correct predictions across 1,2 and 3. This means that our model improves over model 1. This is good since it validates that the preference influence we observed in the deltameans for STC consumers corresponds with higher performance in our model predictions.

```
          ydatavec
custchoice    1    2    3
         1 3640  478  219
         2  256 3661  236
         3  308  450 6016
```

Multiclass Specificity vs. Sensitivity Plots

Relative to model 1, we see a small increase in the rate of our accuracy, 85.4%. we also see a slight improvement in our AUC score from 90.04% to 90.05%.



Additional Scenarios

For the two additional scenarios laid out by Obee, we calculate the overall betameans and then use the model to predict the extra scenario.

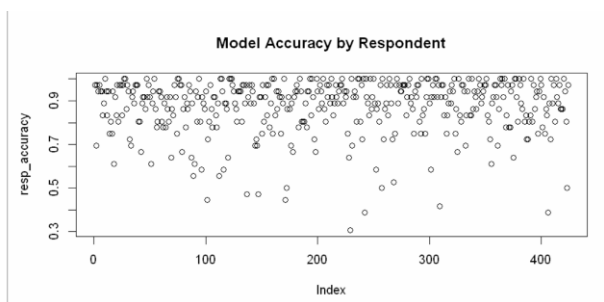
A matrix: 2 × 3 of type dbl

0.07428513	0.6899186	0.23579630
0.51824813	0.4790774	0.00267449

In the first scenario, we see that the STC brand is 68.99% preferred over both Pear and Gaggle with lower configurations for RAM. This is interesting because it suggests that new buyers prefer STC with only a higher RAM configuration than competitors for $199 against Pear and Gaggle.

In the second scenario, we see that a 5-inch screen with low specs and low prices is slightly preferred over Gaggle, which has more RAM. In relation to Pear, there is almost no interest at a higher price point. This means that our brand is stronger against Gaggle, which is able to make up for lesser configurations.

We see that our model performs quite well using our individual model. The below chart shows the actual vs predicted response based on accuracy from the confusion matrix.



Calculating Attribute Level Partworths and Importance

In order to calculate attribute level partworths, we take the average of the beta coefficients across respondents from betadraw. By calculating this across the respondents, we get partworth coefficients.

In order to calculate importances, it is best to compute importances for respondents individually and then average them rather than computing importances from average utilities. Importance measures are ratio-scaled. When we compute an attribute's importance, it is always relative to the other attributes in the study.

In order to obtain a set of attribute importance values that add up to 100 percent, we calculate percentages from relative ranges by taking the maximum and minimum for each attribute and their part-worth utilities.

Tablet Product Recommendation for STC

In conclusion, our first recommendation to STC is to go after the $199 customers. At this price point, Obee's first scenario suggests that we can win a significant amount of preference simply by offering the highest RAM set against competitors. Therefore, we should go with 32 GB RAM, 2 GHZ RAM, and 10-inch screens.

On the other hand, we know that we have power over other brands by offering the same features. A second higher-end offering may help to command even more preference. One thing we may want to avoid though is making our $199 offer look cheap relative to another product that is priced similarly. Consequently, we could offer price matching anywhere from $299 until $399 on configurations put out by our competitors. This would allow us to still command the lower end of the market while taking advantage of our brand strength in the upper price ranges.

One thing to address is the negative perception of STC from existing owners. Existing owners prefer Pear that should be addressed for the purpose of customer lifetime value. These consumers like smaller screens, but they really want additional GHZ in their product. One thing that may make sense is to offer a free upgrade to 2.5 GHZ for a processor for existing customers if they purchase another STC product. In terms of pricing, STC should try and undercut Pear given the lower preference compared to Gaggle and Somesong.

In terms of limitations, the model suggests that the feature offerings are equal across competitors. It could be the case that some brand recognition over certain types of products may be influencing the decision of respondents. In addition, pricing could substantially deviate from the closed-form simulation that we did. There may be factors including the fact that we could raise prices up until $299 or up until $399 without the expected effects from our conjoint model. Scenarios like this could potentially be due to prices that are odd and end in 99 have a different effect on consumer purchase behavior than other prices. In the end, we know that our biggest competitor is Pear, and we know that we can command the $199 price point by offering a feature-rich configuration.