Summary

1 - Homogeneous Logit

About the construction of the data

First, let's define a brand. I've assigned each UPC into a ketchup brand according to the description variable. Based on my understanding of kchp_mksh1.doc there are only 6 brands in market 1. Those are:

- HEINZ
- HUNT'S
- DEL MONTE
- CTL BR
- GENERIC

I identified the brand of each UPC by checking whether the desc variable in the UPC file contained one of the words in the list above. To simplify matters, I considered the CTL BR and the GENERIC to be the store brand STORE.

Second, let's define a purchase occassion by the day-store. If in the household purchase data we observe a household purchasing ketchup more than once in the same store-day I keep only the first purchase.

Third, we need to obtain a price that varies by the household, the brand, and the purchase occasion. Note that households don't get utility from the quantity (weight) of ketchup that they consume, and they dislike paying higher a price. As products have a different size, and fixing a brand, the higher the quantity the higher the price (and presumably the lower the price per ounce) we have to normalize the price paid to represent always the same quantity.

1. I merge the UPC file and the price file, and I construct a panel that contains the price (dpr) for each UPC on each store-day and the weight of each UPC in ounces, that is $weight = \frac{wamt}{1000}$.
2. For each store-day-UPC I compute the price per standarized quantity, which I am going to take as 32 ounces, as it is the most purchased size in the data. That is I compute price as $32 * \frac{dpr}{weight}$.
3. Now, to aggregate from the store-day-UPC to the brand-day-UPC I take the weighted average of the price where the weights vary by the UPC and is given by the total number of standarized units (where the standard is again 32 ounces) sold in the entire panel.
4. At this point, we have a price index that makes sense but it does not vary by the household. To achieve that I will use the data on coupons, and I will assume that a household only has coupons on the product it purchases. The final price will be equal to the one computed in step **3.** minus the total standarized coupons that the household had for that brand the day of the purchase, were we allow for both store and manufacturer couopons. We define standarized coupons as $\frac{32}{weight*units} * (scval + mcval))$, where the division by units is important to normalize the discount to the same 32 ounce bottle unit as the price. An example to illustrate this step is probably a good idea. Let's say that a household on a given day-store had a \$1 coupon when buying a \$2 bottle of ketchup of 32 ounces of a given brand. Then the price faced by the household is \$1. If instead the household bought a 44 ounces bottle of a ketchup brand that costs \$3, then the standarized pre-coupon-price of the bottle would be $32 * (\frac{3}{44}) = 2.18$ and the standardized value of the coupons would be $32 * \frac{1}{44} = 0.73$, so the price faced by the household would be \$1.45. If instead the household bought two 32 ounces bottles of ketchup that cost \$2 each and had a coupon of \$1, then the pre-coupon-price faced would be \$2 (as that is the price for a 32 ounces bottle) and the standardized value of coupons will be \$0.5, so the price faced by the household would be \$1.5.

5. The last step is an assumption. Whenever we observe a household-store-brand with a negative price, we assume it payed zero. There are only 115 observations like these among 26975 purchsaes under consideration.

Fourth, we build a panel as follows. For each household-day-store (which we refer hereafter as a shopping trip) we have 4 observations, one corresponding to each brand. For each shopping trip and brand we also have the corresponding price (as computed by the explanation abive). If for any of the 4 brands in any given shopping trip a price is not available, then we drop the whole shopping trip (the 4 observations, one for each brand). Finally, we add a dummy variable that indicates which of the four brands was chosen for each shopping trip.

(a) Estimate the parameters of the model and comment on whether the parameter estimates make sense to you.

Let's start by defining the probability that a household i at purchase ocassion t choses brand j. That is:

$$\pi_{itj} = Prob(\alpha - \beta P_{itj} + \epsilon_{itj} > \alpha - \beta P_{itk} + \epsilon_{itk}, \ \forall \ k \neq j)$$

Following Train (2009), after some algebra and using the fact that $\epsilon_{itk} - \epsilon_{itj}$ is distributed logistic (because ϵ_{ijt} is distributed i.i.d type 1 extreme value), we have that:

$$\pi_{itj} = \frac{\exp(\alpha - \beta P_{itj})}{\sum_k \exp(\alpha - \beta P_{itk})}$$

Then, the probability that household i at purchase ocassion t is observed choosing alternative j is given by:

$$\Pi_j (\pi_{itj})^{y_{itj}}$$

Where $y_{itj} = 1$ if household i at purchase ocassion t chooses brand j and 0 otherwise. Now assuming that:

1. At any given t the choice of decision maker i is independent of the choice of any other decision maker at any t (including that same t).
2. For any given i the choices over t are independent.

We can write the likelihood of observing a sequence of decision's in the data as:

$$L(\alpha, \beta) = \Pi_{i=1}^N \Pi_{t=1}^{T_i} \Pi_{j=1}^J (\pi_{itj})^{y_{itj}}$$

Therefore, applying the natural logarithm we have that:

$$LL(\alpha, \beta) = \sum_{i=1}^N \sum_{t=1}^{T_i} \sum_{j=1}^J y_{itj} \Big[(\alpha - \beta P_{itj}) - \ln(\sum_k \exp(\alpha - \beta P_{itk})) \Big]$$

To obtain the ML estimates we first normalize $\alpha = 0$ and then we maximize $LL(0, \beta)$ with respect to β.

(b) Now re-estimate supposing that:

$$u_{itj} = \alpha_j - \beta P_{itj} + \epsilon_{itj}$$

and comment on whether the parameter estimates make sense to you. The remaining problem is based on your estimates in this part.

(c) What is the average willingness-to-pay for Heinz ketchup relative to store-brand ketchup?

The average willingness-to-pay for Heinz relative to store-brand, which I will denote as ω, can be computed by equating solving ω from:

$$E[\alpha_{HEINZ} - \beta(P_{itSTORE} + \omega) + \epsilon_{itHEINZ}] = E[\alpha_{STORE} - \beta(P_{itSTORE}) + \epsilon_{itSTORE}]$$

Rearranging using the linearity of the expectation operator:

$$E[\alpha_{HEINZ} - \alpha_{STORE} - \beta\omega + (\epsilon_{itHEINZ} - \epsilon_{itSTORE})] = 0$$

Which we can write as:

$$\alpha_{HEINZ} - \alpha_{STORE} - \beta\omega + E[\epsilon_{itHEINZ} - \epsilon_{itSTORE}] = 0$$

But note that because of the i.i.d. type 1 extreme value distibution of the ϵ_{itj} we know that $E[\epsilon_{itHEINZ} - \epsilon_{itSTORE}] = 0$. Therefore, we have that:

$$\omega = \frac{alpha_{HEINZ} - \alpha_{STORE}}{\beta}$$

(d) What is the expected loss in consumer surplus (per purchase occasion) from eliminating the Del Monte brand from the market?

Start by noting that in this context the partial derivative of utility with respect to income Y_{itj} is such that:

$$\frac{\partial u_{itj}}{\partial Y_{itj}} = \frac{\partial u_{itj}}{\partial(-P_{itj})} = \beta$$

So that, following Train (2009), we have that:

$$E(\Delta CS) = \frac{1}{\beta}\left[\ln\left(\sum_j \exp(\alpha_j - \beta P_{itj})\right) - \ln\left(\sum_{j \neq Del\ Monte} \exp(\alpha_j - \beta P_{itj})\right)\right]$$

(e) What is the elasticity of demand for each brand?

Again, following Train (2009), we have that the price elasticity of demand for each brand is given by:

$$\eta_{it}^j = \frac{\partial \pi_{itj}}{\partial P_{itj}}\frac{P_{itj}}{\pi_{itj}} = -\beta P_{itj}(1 - \pi_{itj})$$

3

(f) Assume that ketchup is produced with constant marginal cost and that each brand chooses its price to maximize its profit, taking as given the prices of the other brands. Given the elasticity in (e) and the average price of each brand, what is the implied marginal cost of each brand?

Assuming that the marginal cost of ketchup is constant, we can write the static (that's why to simplify notation we only index by j) profit function of a brand as:

$$\Pi(P_j, P_{-j}) = (P_j - c_j)\pi_j(P_j, P_{-j})$$

Where P_j is the price of brand j, P_{-j} represents the prices of the other brands, π_j is the logistic probability of choosing alternative j as computed in part **(b)**, and c_j is the marginal cost of brand j.

The first order condition for profit maximization is:

$$\frac{\partial \Pi(P_j, P_{-j})}{\partial P_j} = \pi_j(P_j, P_{-j}) + (P_j - c_j)\frac{\partial \pi_j(P_j, P_{-j})}{\partial P_j} = 0$$

Now, note that:

$$\frac{\partial \pi_j(P_j, P_{-j})}{\partial P_j} = \frac{\pi_j(P_j, P_{-j})}{P_j}\eta^j$$

Where η^j is the price elasticity of demand of brand j. The, we get that the first order condition for profit maximization is:

$$\frac{\partial \Pi(P_j, P_{-j})}{\partial P_j} = \pi_j(P_j, P_{-j}) + \frac{(P_j - c_j)}{P_j}\pi_j(P_j, P_{-j})\eta^j = 0$$

Rearranging we get that:

$$c_j = P_j\left(1 + \frac{1}{\eta^j}\right)$$

(g) Now estimate a more flexible version of the model with:

$$u_{ijt}^* = \alpha_j^* - \beta^* p_{ijt} + \sigma\epsilon_{ijt}^*$$

where σ controls the dispersion in preferences. Does this added flexibility lead to more sensible results?

Note that the σ is not indexed by i or t. Let's define $\tilde{u}_{itj} = \frac{u_{itj}^*}{\sigma}$. Then, we can think of the new model as:

$$\tilde{u}_{itj} = \frac{\alpha_j^*}{\sigma} + \frac{\beta^*}{\sigma}P_{itj} + \epsilon_{itj}$$

But note that this is just the model estimated in **(b)** where:

$$\alpha_j = \frac{\alpha_j^*}{\sigma} \quad \beta = \frac{\beta^*}{\sigma}$$

Therefore, this does not add more flexibility than before.

4

(h) Return to the model in (2) and suppose that the cost estimates in (f) are correct. Suppose now that Del Monte exits the market and that the remaining brands choose new prices simultaneously. Assuming that the new prices constitute a Nash equilibrium of the price-setting game, what is the resulting loss in consumer surplus?

If Del Monte exits the market, we have that for any remaining brand j:

$$\tilde{\pi}_j(P_j, P_{-j}) = \frac{\exp(\alpha_j - \beta P_j)}{\sum_k \exp(\alpha_k - \beta P_k)}$$

Therefore, a remaining brand j trying to maximize profits taking as given the prices of the other remaining brands will solve:

$$argmax_{P_j} \ (P_j - c_j)\tilde{\pi}_j(P_j, P_{-j})$$

After some algebra, the first order condition of this problem is:

$$[(P_j - c_j)\beta - 1]\frac{\sum_{k \neq j} \exp(\alpha_k - \beta P_k)}{\exp(\alpha_j - \beta P_j)} = 0$$

Which defines implicitly P_j^{BR} for a remaining brand j. For simplicity, assume that Del Monte was indexed by brand 4. The new prices $(P_1^{NEW}, P_2^{NEW}, P_3^{NEW})$ constitute a Nash equilibrium if they are a fixed point of the system of equations given by the 3 first order conditions of the remaining firms. We can solve the stated system numerically.