FEB23002 Marketing Models
Computer Assignment
Academic Year 2021-2022

Note this assignment contains 2 pages.

Instructions

This computer exercise is part of the exam for the course Marketing Models. The rules for this assignment are given on slide 4 of week 1. It is an individual assignment. You can use any software package you like to make this exercise. Hand in via CANVAS before the deadline:

- The answering sheet with your answers (filled-in pdf file)

Exercise

On Canvas there are 100 data sets containing brand and size choice decisions on purchasing tomato ketchup. The data set you have to analyze is determined by the last 2 digits of your student number. Each dataset is a random sample of a larger dataset containing real purchase decisions.

Households only buy 1 brand per store visit and there are 4 brands-size combinations available (Heinz 28 ounces, Heinz 32 ounces, Heinz 40 ounces, Hunts 32 ounces). You have in total 2,000 store visits, where ketchup is purchased. Per brand choice decision you have the brand-size choice (4 binary dummies), the price of each brand-size, a binary feature promotion variable for each brand-size, and a binary display promotion variable for each brand-size. You have access to data in two formats. The first format contains contain one row for each choice situation (purchase decision). The second format contains 4 rows per choice situation (purchase decision - the 4 rows correspond to the 4 brand-size alternatives).

The original dataset contained multiple observations per household. For the purpose of this exercise, the household ID variable has been removed. Please analyse the data as if all observations are from separate individuals. i.e. do not adjust the analysis to account for multiple observations for some households (this is beyond the course syllabus and it is not possible anyway without the household ID variable).

1. Consider a Conditional Logit model to describe the brand choice between the 4 ketchup brand-sizes, where you include as explanatory variables intercepts and price.

Set the intercept specific to Heinz 28 ounces to zero for identification. Estimate the parameters of this model using Maximum Likelihood (you can use any package you want) and report the parameter estimates together with regular standard errors in the answering sheet.

2. Report the maximum log likelihood value and the McFadden R^2.

3. The partial/marginal (cross-)price effects are defined as

$$\frac{\partial \Pr[Y = j]}{\partial \text{price}_k},$$

where Y denotes the random multinomial choice variable with $j \in \{\text{Heinz 28 ounces, Heinz 32 ounces, Heinz 40 ounces, Hunts 32 ounces}\}$ and price_k denotes the price of brand $k \in \{\text{Heinz 28 ounces, Heinz 32 ounces, Heinz 40 ounces, Hunts 32 ounces}\}$.

Compute the average partial effects of the prices of each brand-size by averaging the partial/marginal effects over the observations in the sample, that is,

$$\frac{1}{N} \sum_{i=1}^{N} \frac{\partial \Pr[Y_i = j]}{\partial \text{price}_{ik}},$$

where $N = 2,000$ is your sample size. Report these average marginal/partial effects in the table in the answering sheet.

4. What are the average effects of a display promotion for Hunts 32 ounces (with no feature promotion for Hunts and no feature nor display promotion for any other brand-sizes) on the probabilities of all 4 brand-sizes relative to a scenario in which there are no feature nor display promotions for any brand-sizes (including Hunts). [Here "average effects" means differences in estimated probabilities averaged over the sample]

5. Test the assumption of the independence of irrelevant alternatives by implementing a Hausman test comparing the unrestricted model with a restricted model that does not include the Heinz 40 ounces observations. Report the test statistic and p-value in the spaces provided in the answer sheet.