Assignment 1: Linear Demand Estimation

Econ 543

Due: September 20

TA Session: September 27

This homework assignment will give you practice working with the basics of the multinomial logit and nested logit demand models. It is closely related to Berry (1994) but is limited in terms of consumer heterogeneity. The following homework (problem set 2) will introduce simulation-based estimators to provide much richer sources of consumer heterogeneity.

The file **product_data_2.csv** contains information on 10 different soft drinks, for 100 different markets (or time periods). This dataset contains the following variables: product_ID (the product ID), nest (a variable indicating if the product is a diet or a regular soda), price (the product's price during period t, in USD), **quantity (the number of cans sold by each product during period t)**, sugar (the grams of sugar contained in product j at time t), caffeine (the milligrams of caffeine contained in product j at time t), caffeine_extract_price (the price of caffeine extract paid by the firm manufacturing product j at time t), corn_syrup_price (the price of high fructose corn syrup paid by the firm manufacturing product j at time t), and t (the observation's time period). **Moreover, we know that, in all markets, there are $\bar{M}_t = 100,000$ consumers.**

You may work together on this assignment but every student should submit their own responses in LaTeX, markdown or similar in addition to code via canvas by the due date. We will go over the assignment during the TA session.

1 Multinomial Logit

Problem 1(a): Multinomial Logit. Assume the indirect utility of consumer i for consuming soft drink j in period t is given by:

$$u_{ijt} = \alpha p_{jt} + \beta_1 sug_{jt} + \beta_2 caf_{jt} + \gamma_d Diet_j + \gamma_r Regular_j + \xi_{jt} + \varepsilon_{ijt} \tag{1}$$

where p_{jt} denotes the price (in USD) of product j at time t, sug_{jt} denotes j's sugar content (in grams) at time t, caf_{jt} denotes j's milligrams of caffeine at time t, $Diet_j$ is an indicator variable equal to 1 if j is a diet product, $Regular_j$ is an indicator variable equal to 1 if j is a regular soda, and ξ_{jt} represents the unobserved quality of product j during time period t.

Moreover, assume that ε_{ijt} are *iid* draws from a standard type 1 extreme value distribution. Note that the only source of heterogeneity between consumer preferences comes from the extreme value shock.

Finally, assume that there is an outside option (product 0) representing not consuming any soft drinks. The utility of the outside option is normalized such that it is mean zero:

$$u_{i0t} = \varepsilon_{i0t} \tag{2}$$

Recall that given our distributional assumptions over ε_{ijt}, the probability that consumer i chooses product j during time t is:

$$s_{ijt} = \mathbb{P}(u_{ijt} > u_{ikt} \text{ for all } k) = \frac{\exp(\delta_{jt})}{\sum_k \exp(\delta_k) + 1}$$

With $\delta_{jt} = \alpha p_{jt} + \beta_1 sug_{jt} + \beta_2 caf_{jt} + \gamma_d Diet_j + \gamma_r Regular_j + \xi_{jt}$.

Assume that p_{jt}, sug_{jt}, and caf_{jt} are all exogenous variables. Estimate the parameters of this model (α, β_1, β_2, γ_d, and γ_r).

Problem 1(b). Now assume that sug_{jt}, caf_{jt}, $Diet_j$ and $Regular_j$ are exogenous, but that p_{jt} is correlated with the unobserved quality ξ_{jt}. The caffeine_extract_price$_{jt}$ and corn_syrup_price$_{jt}$ affect prices in proportion to caffeine and sugar content, so use them to construct instruments for prices. Describe your instruments and present your estimates of the parameters of the model. What conditions must these two variables satisfy to be valid instruments?

Problem 1(c). Use the parameters you estimated in problem 1 (b) to calculate the vector of price own-price derivatives $\frac{\partial s_{jt}}{\partial p_{jt}}$ for all j and all t and vector of the own-price elasticities $\frac{\partial s_{jt}}{\partial p_{jt}} \frac{p_{jt}}{s_{jt}}$ for all j and all t. What is the mean own-price elasticity in the data for Regular and Diet drinks, respectively?

Problem 1(d). Use the parameters you estimated in problem 1 (b) to derive the vector of cross-price derivatives $\frac{\partial s_{jt}}{\partial p_{1t}}$ for all $j \neq 1$ and all t. Use this vector to calculate the vector of cross-price elasticities $\frac{\partial s_{jt}}{\partial p_{1t}} \frac{p_{1t}}{s_{jt}}$ for all $j \neq 1$ and all t. What is the mean cross-price elasticity between product 1 and diet sodas? What about between product 1 and regular sodas?

Problem 1(e). Write a function that generates the Jacobian matrix of price derivatives $\Delta(p)$ (i.e., $\Delta_{ij} = \frac{\partial s_j}{\partial p_i}$) for a given period. Your function should take the time period, product variables, and model parameters as arguments, and return the Jacobian matrix. Print the Jacobian for the last time period in the sample (t = 100).

2 Nested Logit

Problem 2(a): Nested Logit. Now suppose that there are two sources of consumer heterogeneity (1) an extreme value shock as in problem 1 and (2) a preference for diet soda. Let products $0, 1, 2, ..., 10$ be grouped into three mutually exclusive sets: $\mathcal{J}_D$ (for diet products), $\mathcal{J}_R$ (for regular products), and $\mathcal{J}_0$ (for the outside option). For product $j \in \mathcal{J}_g$ during time t, let consumer i derive utility equal to:

$$u_{ijt} = \alpha p_{jt} + \beta_1 sug_{jt} + \beta_2 caf_{jt} + \gamma_d Diet_j + \gamma_r Regular_j + \xi_{jt} + \varsigma_{igt} + (1-\sigma) \times \varepsilon_{ijt} \tag{3}$$

Where ε_{ijt} is again distributed type 1 extreme value, and ς_{igt} follows a distribution such that $\varsigma_{igt} + (1 - \sigma) \times \varepsilon_{ijt}$ is also an extreme value random variable. The parameter σ represents the degree of correlation between the error terms of products in the same nest and satisfies $0 \leq \sigma < 1$.

Given our distributional assumptions, the inside share of good j belonging to group g at time t is given by:

$$s_{jt|g} = \frac{\exp(\frac{\delta_{jt}}{1-\sigma})}{\sum_{k \in \mathcal{J}_g} \exp(\frac{\delta_{kt}}{1-\sigma})} \tag{4}$$

with $\delta_{jt} = \alpha p_{jt} + \beta_1 sug_{jt} + \beta_2 caf_{jt} + \gamma_d Diet_j + \gamma_r Regular_j + \xi_{jt}$

The probability that a product from group g is chosen is

$$s_g = \frac{D_g^{(1-\sigma)}}{\sum_g D_g^{(1-\sigma)}} \tag{5}$$

with $D_g = \sum_{k \in \mathcal{J}_g} \exp(\frac{\delta_{kt}}{1-\sigma})$

Therefore, product j's share at time t is

$$s_{jt} = s_{jt|g} \times s_g \tag{6}$$

Estimate the parameters of the model (α, β_1, β_2, γ_d, γ_r, and σ), assuming as above that price is correlated with unobserved quality. Write the estimating equation for this model. What instruments are needed to estimate it? What are the parameter estimates?

Problem 2(b). Use the parameters you estimated in problem 2 (a) to calculate the vector of price own-price derivatives $\frac{\partial s_{jt}}{\partial p_{jt}}$ for all j and all t and vector of the own-price elasticities $\frac{\partial s_{jt}}{\partial p_{jt}} \frac{p_{jt}}{s_{jt}}$ for all j and all t. What is the mean own-price elasticity in the data for Regular and Diet drinks, respectively?

Problem 2(c). Use the parameters you estimated in problem 2 (a) to derive the vector of cross-price derivatives $\frac{\partial s_{jt}}{\partial p_{1t}}$ for all $j \neq 1$ and all t. Use this vector to calculate the vector of cross-price elasticities $\frac{\partial s_{jt}}{\partial p_{1t}} \frac{p_{1t}}{s_{jt}}$ for all $j \neq 1$ and all t. What is the mean cross-price elasticity between product 1 and diet sodas? What about between product 1 and regular sodas?

Problem 2(d). How do the results you obtained from parts 2 (a), 2 (b), and 2 (c) compare to the ones from parts 1 (b), 1 (c), and 1(d)?

Problem 2(e). Write a function that generates the Jacobian matrix of price derivatives $\Delta(p)$ (i.e., $\Delta_{ij} = \frac{\partial s_j}{\partial p_i}$) for a given period, under the nested logit specification. Your function should take the time period, product variables, and model parameters as arguments, and return the Jacobian matrix. Print the Jacobian for the last time period in the sample (t = 100).

3 A simple supply side

Problem 3(a). Assume that the supply side is a single-product Bertrand Nash Equilibrium (each firm owns only one product). That is, firm f owning product j solves the following problem in period t to set prices:

$$\max_{p_{jt}} \Pi_{ft} = \max_{p_{jt}} (p_{jt} - c_{jt}) q_{jt}$$

Where c_{jt} denotes product j's marginal cost, **and** $q_{jt} = s_{jt}\bar{M}_t$**.** $\bar{M}_t = 100,000$ **is the number of consumers in this market.** Use the firms' FOCs and your parameters above to estimate the marginal cost for all j and all t. What is the mean Lerner Index for these products?

Problem 3(b). Now assume that the firms who own products 1 and 2 merge at $t = 100$, so that a single firm now owns both products. However, costs $c_{1,100}$ and $c_{2,100}$ are unaffected by the merger. Simulate how the prices of all firms would change in this scenario.

Problem 3(c). How does the average price of the products from the merging firms change after the merger? How does the average price of competing products in the same nest change after the merger? How does the average price for competing products in a different nest change after the merger?

Problem 3(d). Now instead assume that all firms collude during the last period (t = 100) and choose the prices that maximize the sum of all profits. Simulate how the prices of all firms would change in this scenario. Summarize the results of 3(c) and 3(d) in a simple table.