Market Analysis with Econometrics and Machine Learning

3b Discrete Choice Models II: Substitution Patterns and Mixed Logit Models

Uni Ulm

Prof. Dr. Sebastian Kranz

SoSe 2020

Profit maximization for a multi-product firm

- Consider J products with prices $p = (p_1, \ldots, p_J)$.
- The demand $q_j(p)$ for each product j can in principle depend on the prices of all other products.
- Consider a firm who owns products 1 and 2 and thus has profit function

$$\pi(p) = q_1(p)(p_1 - c_1) + q_2(p)(p_2 - c_2)$$

- The first order conditions for the profit maximizing prices p_1 and p_2 are

$$\frac{\partial \pi(p)}{\partial p_1} = q_1(p) + \frac{\partial q_1(p)}{\partial p_1}(p_1 - c_1) + \frac{\partial q_2(p)}{\partial p_1}(p_2 - c_2) = 0$$

$$\frac{\partial \pi(p)}{\partial p_2} = q_2(p) + \frac{\partial q_1(p)}{\partial p_2}(p_1 - c_1) + \frac{\partial q_2(p)}{\partial p_2}(p_2 - c_2) = 0$$

Interpreting the first order conditions

- Consider the FOC with respect to the price of product p_1:

$$\frac{\partial \pi(p)}{\partial p_1} = q_1(p) + \frac{\partial q_1(p)}{\partial p_1}(p_1 - c_1) + \frac{\partial q_2(p)}{\partial p_1}(p_2 - c_2) = 0$$

- The derivative $\frac{\partial \pi(p)}{\partial p_1}$ measures approximately by how much the total profits of our firm increase, if we increase p_1 by 1 Euro.
- The first term $q_1(p)$ describes the additional profit that we make on all sales of product 1 because the price increases by 1 Euro.
- The second term $\frac{\partial q_1(p)}{\partial p_1}(p_1 - c_1)$ will typically be negative and measures how much profit we forgo because the price increase reduces the demand of product 1. The *own-price derivative* of demand $\frac{\partial q_1(p)}{\partial p_1}$ approximates by how many units sales of product 1 go down if p_1 increases by one Euro and the markup $p_1 - c_1$ is the profit per unit sold.
- The third term $\frac{\partial q_2(p)}{\partial p_1}(p_2 - c_2)$ measures the impact on our profits due to the change in the demand of the second product. The *cross price derivative* of demand $\frac{\partial q_2(p)}{\partial p_1}$ approximates the effect of a one Euro increase in the price of product 1 on the demand for product 2.
    - If the two products are *substitutes* (e.g. two different car models), the cross price derivative will be positive. A price increase of product 1 increases demand for product 2.
    - If the two products are *complements* (say a video game console and a video game for it), the cross price derivative is negative and a price increase of product 1 reduces profits for product 2.

Price Elasticities

- Empirical studies often report *price elasticities*, which approximate by how much *percent* demand changes if a price changes by one *percent*.
- The *own price elasticity* of a product j is given by

$$\eta_{jj} = \frac{\partial q_j}{\partial p_j}\frac{p_j}{q_j}$$

- The *cross price elasticity* of a product j with respect to a price change of product $h \neq j$ is given by

$$\eta_{jh} = \frac{\partial q_j}{\partial p_h}\frac{p_h}{q_j}$$

- Since price elasticities measure percentage changes, they are much easier to compare between products than price derivatives of a demand function.

Discrete Choice Models to Estimate Demand Functions for Differentiated Products

- By estimating utility functions of discrete choice models, one can derive demand functions and corresponding price derivatives and elasticities for *differentiated products*, i.e. (imperfect) substitutes.
- We have already seen in our heating choice application how one can aggregate choice probabilities of many consumers to expected market shares s_j for each product. In general each market share will depend on the prices and other product attributes of all product alternatives.
    - From a choice data set (choice experiment or sample from the real world), we can estimate markets shares but typically don't know the total market size. Here you have to make an educated guess based on other sources.
- If we assume the total market size is M, demand for product j can be written as $q_j = Ms_j$.
    - A price derivative is given by

$$\frac{\partial q_j}{\partial p_h} = M\frac{\partial s_j}{\partial p_h}$$

    - Price elasticities can be computed from the market share functions without knowing the market size:

$$\eta_{jh} = \frac{\partial q_j}{\partial p_h}\frac{p_h}{q_j} = \frac{\partial s_j}{\partial p_h}\frac{p_h}{s_j}$$

How flexible is a conditional logit model?

- If we estimate a conditional logit model, we make a lot of assumptions, e.g. that the unobserved random utility shocks ε_{nj} are indepedently drawn for each combination of consumer n and choice j and that they follow a Gumbel distribution.
- Whenever we make assumptions in an econometric model, we would like to better understand their consequences.
- In particular, do such assumptions impose strong restrictions on the structure of the price derivatives that we can possibly estimate? Is our model able to capture realistic substitution patterns, e.g. that two sports car models may be closer substitutes than a sports car and a family car?
- We will explore this question on the following slides.

- Assume we have a discrete choice data set about car choices without any background characteristics about the consumers.
- Assume a utility function of the form

  $$U_{nj} = -\alpha p_j + \beta_1 x_{1j} + \ldots + \beta_K x_{Kj} + \varepsilon_{njk}$$

  where p_j is the price of the car and the other explanatory variabels x_{kj} describe some product characteristic like horse power, or fuel efficiency.
- Note that the explanatory variables only contain product attributes, no consumer characteristics. So consumers only differ by their random utility shock ε_{njk} and the deterministic part of the utility function is the same for all consumers:

  $$V_{nj} = -\alpha p_j + \beta_1 x_{1j} + \ldots + \beta_K x_{Kj} \equiv V_j, \forall n$$

Price derivatives of market shares in conditional logit model without consumer characteristics

- The expected market share of product j in such a conditional logit model is simply its choice probability:

$$s_j = P_j = \frac{exp(V_j)}{\sum_{h=1}^{J} exp(V_h)}$$

- We can show that the own-price derivative of the market share for product j in this conditional logit model is given by

$$\frac{\partial s_j}{\partial p_j} = -\alpha s_j(1 - s_j)$$

  and the cross price derivative of the market share for product j with respect to the price of product $h \neq j$ is given by

$$\frac{\partial s_j}{\partial p_h} = \alpha s_j s_h$$

Optional: Derivation

We will derive the formula for the own price derivative of a product's market share. Using the chain rule of differentiation, we have

$$\frac{\partial s_j}{\partial p_j} = \frac{\partial s_j}{\partial V_j}\frac{\partial V_i}{\partial p_j}$$

We have $\frac{\partial V_i}{\partial p_j} = -\alpha$. Let us now compute the derivative $\frac{\partial s_j}{\partial V_j}$. Assuming we already normalized the scale parameter of ε_{nj} to $\sigma = 1$, we have

$$s_j = \frac{\exp(V_j)}{\sum_{h=1}^{J}\exp(V_h)} \equiv \frac{N}{D}$$

where $N = \exp(V_j)$ and $D = \sum_{h=1}^{J}\exp(V_h)$ shall be the numerator and denominator of this ratio. The quotient rule of differentiation tells us that

$$\frac{\partial s_j}{\partial p_j} = \frac{\frac{\partial N}{\partial p_j}D - N\frac{\partial D}{\partial p_j}}{D^2}.$$

We have

$$\frac{\partial N}{\partial p_j} = \frac{\partial \exp(V_j)}{\partial p_j} = \frac{\partial V_j}{\partial p_j}\frac{\partial \exp(V_j)}{\partial V_j} = -\alpha\exp(V_j) = -\alpha N$$

and

$$\frac{\partial D}{\partial p_j} = \frac{\partial \sum_{h=1}^{J}\exp(V_h)}{\partial p_j} = \frac{\partial \exp(V_j)}{\partial p_j} = -\alpha\exp(V_j) = -\alpha N$$

Hence we have

$$\frac{\partial s_j}{\partial p_j} = \frac{-\alpha ND + \alpha N^2}{D^2} = -\alpha(\frac{N}{D} - \frac{N^2}{D^2}) = -\alpha\frac{N}{D}(1 - \frac{N}{D}) = -\alpha s_j(1 - s_j)$$

.

The formula for the cross price derivative $\frac{\partial s_j}{\partial p_h}$ can be derived in a similar fashion.

- The cross price elasticities for $h \neq j$ are given by

  $$\frac{\partial s_j}{\partial p_h}\frac{p_h}{s_j} = \alpha s_h p_h$$

  They only depend on the product whose price is changed.

- The cross price derivative

$$\frac{\partial s_j}{\partial p_h} = \alpha s_j s_h$$

  in our model only depends on the marginal effect α of prices on utility and on the market share of both products.
- Consider for example 3 cars: a sports car A (say a Mercedes SLK), a similar sports car B (say a Porsche Boxter) and a family car C (say a Volkswagen Passat). We estimate the following market shares: $s_A = 1\%$, $s_B = 1\%$ and $s_C = 5\%$, we have a market size M of 1 Million and assume that prices are denoted in 1000 Euros and we have $\alpha = 1$. Assume the price of sports car A increases by a 1000 Euros. Has this a stronger effect on the demand for the other sports car B or for the family car C?
- In our conditional logit model without consumer characteristics we find the following cross price derivatives:

$$\frac{\partial q_B}{\partial p_A} = \alpha M s_A s_B = 100$$

$$\frac{\partial q_C}{\partial p_A} = \alpha M s_A s_C = 500$$

- This means we would predict that a 1000 Euro price increase of sports car A, increases the sales of sportscar B by 100 cars and that of the family car C by 500 cars: 5 times the amount.
- We just predict that the effect on the family's car demand is 5 times as large because it has originally 5 times as large a market share than the other sports car B.
- But this prediction is not very realistic. In the real worls we would expect that the two sports cars are closer substitutes, i.e. a price increase of sports car A should have a stronger effect on the sales of the similar sports car B than on the quite different family car C.
- This limited substitution pattern that only takes into account the market shares of two products is a severe limitation of our conditional logit model without consumer characteristics.

The red bus - blue bus problem

- Another illustration of the limitations of our conditional logit model without consumer characteristics is the red bus - blue bus problem.
- Assume consumers can initially only choose between taking a car to work $j = 1$ or a red bus $j = 2$ and assume both alternatives have the same deterministic utility $V_1 = V_2 = 1$. Thus both have the same market share $s_1 = s_2 = \frac{1}{2}$.
- Assume now a blue bus is introduced as a third alternative that, except for its color which consumers don't care about, is absolutely identical to the red bus. So we we have $V_3 = V_2 = 1$.
- What will now be the new market shares in a conditional logit model without consumer characteristics?
- Since all 3 alternatives give consumers the same deterministic utility you quickly find with the formula for choice probabilities that $s_1 = s_2 = s_3 = 1/3$, i.e. the total market share of all buses is now 2/3.
- Realistically, if both buses are identical expcet for the color (and there was no crowding before) we would expect however that the market share of cars stays close to 1/2 and each bus has a market share of roughly 1/4.

Independence of irrelevant alternatives

- The limited substitution patterns can be explained by a feature of the conditional logit model that is called *independence of irrelevant alternatives*.
- It means that that the ratio of the choice probabilities of two alternatives is independent of the presence or features of any other alternatives:

$$\frac{P_{nj}}{P_{nh}} = \frac{\exp(V_{nj})}{\exp(V_{nh})}$$

- In a conditional logit model without consumer characteristic, this implies that ratio of expected market shares of two products is independent of other alternatives:

$$\frac{s_j}{s_h} = \frac{\exp(V_j)}{\exp(V_h)}$$

- Assume we have 4 products: Products A and B are almost identical (say two very similar sports cars) and products C and D are almost identical (say two very similar family cars).
- Each of the products shall have the same expected market share $s_A = s_B = s_C = s_D = 0.25$
  - This means a conditional logit model without consumer characteristics would predict that a price increase by product A has the same effect on product B, C and D.
- Assume we have two equally sized consumer groups $i = 1, 2$ one prefering sports cars and the other family cars with approximatly the following expected market shares s_{ij} within each consumer group.

Approximate market shares	A	B	C	D
Consumer group 1	0.5	0.5	0	0
Consumer group 2	0	0	0.5	0.5
Total market	0.25	0.25	0.25	0.25

- Within each consumer group demand shall be characterized by a conditional logit model without any additional consumer characteristics playing a a role (and a linear effect α on price)

- The total market share for product j is thus the weighted sum of the market share in all consumer groups. If $w_i = 1/2$ is the population share of consumer group i, we have:

$$s_j = \sum_{i=1}^{2} w_i s_{ij} = \frac{1}{2} s_{1j} + \frac{1}{2} s_{2j}$$

- A cross price derivative of market shares is now given by

$$\frac{\partial s_j}{\partial p_h} = \sum_{i=1}^{2} w_i \frac{\partial s_{ij}}{\partial p_h} = \frac{1}{2}\frac{\partial s_{1j}}{\partial p_h} + \frac{1}{2}\frac{\partial s_{2j}}{\partial p_h} = \frac{1}{2}\alpha s_{1j} s_{1h} + \frac{1}{2}\alpha s_{2j} s_{2h}$$

- We find a cross price derivative between the two sports cars of

$$\frac{\partial s_B}{\partial p_A} = \frac{1}{2}\alpha(s_{1B}s_{1A} + s_{2B}s_{2A}) = \frac{1}{2}\alpha(\frac{1}{2}\cdot\frac{1}{2} + 0 \cdot 0) = \frac{\alpha}{8}$$

  and between the sport and familiy car

$$\frac{\partial s_C}{\partial p_A} = \frac{1}{2}\alpha(s_{1C}s_{1A} + s_{2C}s_{2A}) = \frac{1}{2}\alpha(0 \cdot \frac{1}{2} + \frac{1}{2} \cdot 0) = 0$$

- Even though for each consumer type the cross price derivaties only depend on the market share of a product (and the fixed α), the cross price derivatives of the total market allow more complex substition patterns. In our example, a price increase of sports car A only affects the other sports car B but not the family car C.

Mixed Logit Models

- Mixed logit models are popular discrete choice models that allow flexible substitution patterns by assuming there are different unobserved consumer types indexed by i.
- If consumer n is a consumer type i her utility is given by

$$U_{nj} = \beta_{1i} x_{1nj} + \ldots + \beta_{Ki} x_{Knj} + \varepsilon_{nj}$$

  - β_{ki} measures how much utility of a consumer type i increases if attribute x_{knj} increases by one unit.
- So in addition to the random utility shock ε_{nj}, we now have unobserved heterogeneity by the degree of how much consumers value different product attributes. This allows for flexible substitution patterns.
- In our car example above, imagine that we have dummy variables that indicate family cars and sports car. Consumer type 1 had a very high β_{k1} for the family car dummy and consumer type 2 for the sports car dummy.
- The model is called *mixed logit* because expected choice probabilities for a product are a mix of the choice probabilities for every consumer type.

Modelling the distribution of β_{ki}

- In a mixed logit model one assumes that each coefficient β_{ki} follows a distribution, e.g. a normal distribution with some unknown mean $\bar{\beta}_k$ and standard deviation σ_k.
  - These *meta parameters* $\bar{\beta}_k$ and σ_k will be estimated when estimating a mixed logit model.
- Since we treat the coefficients β_{ki} as randomly drawn for each consumers they are also called *random coefficients*.
- One can specify different distributions for different coefficients, assume that some coefficients are fixed for all consumers, or allow correlations between the coefficients β_{ki}.
- We will explore the mixed logit model in the RTutor problem set 3b for a simulated data set of a fictious discrete choice experiment to understand demand for cars.