Chapter 5 *Discrete choice random utility models*

Based on Regis Renault's Slides

Andras Niedermayer[1]

[1]Department of Economics (THEMA), Cergy Paris Université

Winter 2025

Outline

1. Background on product differentiation.
2. Linear random utility model of demand.
3. Multinomial logit
4. Oligopoly and price competition.
5. Mixed logit and random coefficients.

Horizontal versus vertical differentiation

- Product's attributes often thought of as summarized in its quality (e.g. a better computer has a faster chip, more memory and better video).
- Then consumers would agree on which product is better.
- However consumer tastes are heterogeneous: some might want a big high performance laptop, while others would prefer a small and light one.
- The theory of imperfect competition with product differentiation distinguishes
  - Vertical product differentiation: with equal prices all consumers would prefer one product to the other.
  - Horizontal differentiation: at equal prices, different consumers would select different products.

Three remarks

- The two types of differentiation have different implications for pricing, both in monopoly and oligopoly..
- There can be taste heterogeneity with vertical differentiation (e.g. the Mussa Rosen setting).
- Empirical challenge in identifying and measuring product quality.

Localized versus non localized competition

- Product choice is an essential dimension of non price competition (Hotelling, 1929).
- With more than two products each firm might compete only with a subset of other firms (localized competition, e.g. the Vickrey/Salop circle model where each competes with its two neighbors).
- With enough dimensions in the product space, a firm can compete with all others (non localized competition).

1 Background on product differentiation
Discrete choice versus multi homing

- Discrete choice means that each consumer consumes only one of the products.
- A consumer could conceivably purchase multiple products from multiple sellers (situation described as "multi homing" in the platform literature).
- If this is the case then there is some complementarity between products which softens competition.
- Early models of non localized competition assumed a representative consumer that consumes all varieties, rather than discrete choice (Spence, Dixit, Stiglitz). This approach has not been pursued in IO (rather in Trade, macro and growth theory).

2 Linear random utility model of demand

- For each consumer $\ell = 1, ..., L$ there are n draws $\epsilon_{\ell,i}$, $i = 1, ..., n$, where $\epsilon_{\ell,i}$ are i.i.d across consumers ℓ and $E\epsilon_{\ell,i} = 0$ for all ℓ and i.
- Support of $(\epsilon_{\ell,1}, ..., \epsilon_{\ell,n})$ is $[a, b]^n$, density is f.
- Consumer ℓ's utility if she buys product i at price p_i is

$$u_{\ell,i}(p_i) = \alpha_i - p_i + y_\ell + \epsilon_{\ell,i}. \tag{1}$$

  where y_ℓ is ℓ's income and α_i is a mean valuation for product i (it could for instance depend on some vector of product characteristics X_i with $\alpha_i = X_i\beta$).
- Let $v_i = \alpha_i - p_i$.

2 Linear random utility model of demand

- ℓ prefers product i to product j if and only if

$$v_i + \epsilon_{\ell,i} > v_j + \epsilon_{\ell,j}. \tag{2}$$

- Consumer ℓ prefers not to buy if

$$\max_{i=1,\ldots,n} \{v_i + \epsilon_{\ell,i}\} < 0. \tag{3}$$

- No income effect.
- $\alpha_i + \epsilon_{\ell,i}$ can be interpreted as a measure of the match quality between consumer ℓ and product i.

Determination of demand for product i.

- Consider product 1.
- Consumer ℓ buys product 1 if

$$\epsilon_1 > -v_1 \tag{4}$$

and

$$\epsilon_j < \epsilon_1 + v_1 - v_j. \tag{5}$$

Expected demand for product 1 can therefore be written as

$$D_1(p_1, ..., p_n) = L \int_{-v_1}^{b} \int_{a}^{\epsilon_1 + v_1 - v_2}, .., \int_{a}^{\epsilon_1 + v_1 - v_n} f(\epsilon_1, .., \epsilon_n) d\epsilon_n, .., d\epsilon_1 \tag{6}$$

The joint distribution could be normal $(\epsilon_1, ..., \epsilon_n) \sim N((0, ..., 0), \Sigma)$, which yields a multinomial probit model.

With two products product 1's demand is

$$D_1(p_1, p_2) = L \int_{-v_1}^{b} \int_{a}^{\epsilon_1 + v_1 - v_2} f(\epsilon_1, \epsilon_2) d\epsilon_2 d\epsilon_1. \tag{7}$$

The set over which the integral in calculated can be depicted graphically.
This shows how changes in prices affect product 1's market share.

Because of budget constraints the maximum relevant price is $\bar{p} = \max\{y_1, ..., y_n\}$ (none of the products could be purchased by any consumer at larger prices)
Now assume that $a + \min_i\{\alpha_i\} > \bar{p}$, then for any product i, $-v_i = p_i - \alpha_i < a$. .
Then all consumers buy one product with probability one and product 1's demand is.

$$D_1(p_1, ..., p_n) = L \int_a^b \int_a^{\epsilon_1 + v_1 - v_2}, .., \int_a^{\epsilon_1 + v_1 - v_2} f(\epsilon_1, .., \epsilon_n) d\epsilon_n, .., d\epsilon_1. \tag{8}$$

Specification is only useful to buy some tractability in the theory (especially when combined with symmetry or duopoly).

2 Linear random utility model of demand
i.i.d. matches across products

Assume now that $\epsilon_{\ell,i}$, $\ell = 1, ..., L$, $i = 1, ..., n$ are not only i.i.d. across consumers ℓ but also across products.

F and f respectively denote the common c.d.f and density of $\epsilon_{\ell,i}$.

Then product 1's demand is.

$$D_1(p_1, ..., p_n) = L \int_{-v_1}^{b} \Pi_{i=2}^{n} F\left(\epsilon_1 + v_1 - v_i\right) f(\epsilon_1) d\epsilon_1. \qquad (9)$$

Much of existing theoretical work uses variants of this version.

The multinomial logit is a special case.

Special case of previous one where $\alpha_i = \alpha_j = \alpha$ for all $i, j = 1, ..., n$.

If all products have the same price p^* then all products have the same demand

$$D_1(p^*, ..., p^*) = L \int_{p^*-\alpha}^{b} F(\epsilon_1)^{n-1} f(\epsilon_1) d\epsilon_1 = \frac{L}{n} (1 - F(p^* - \alpha)^n) \tag{10}$$

If market is covered, $p^* - \alpha$ replaced by a so each demand is $\frac{L}{n}$.

2 Linear random utility model of demand
Price sensitive individual demand.

Rather than assuming unit demand assume that, if consumer ℓ buys product i, she purchases a quantity $q_{\ell,i}$ and utility is now given by

$$u_{\ell,i}(p_i) = \alpha_i + v(p_i) + y_{\ell,i}, \tag{11}$$

Where $v(p_i) + y_\ell$ is the indirect utility from consumer product i at price p_i, which is strictly decreasing and convex in p_i (here, since there is no income effect, it is merely consumer surplus).
Then at price p_i consumer buys $d_i(p_i) = -v'(p_i)$.
Letting $v_i = \alpha_i + v(p_i)$ demand is

$$D_1(p_1, ..., p_n)$$
$$= Ld_1(p_1) \int_{-v_1}^{b} \int_{a}^{\epsilon_1+v_1-v_2}, ..., \int_{a}^{\epsilon_1+v_1-v_n} f(\epsilon_1, ..., \epsilon_n) d\epsilon_n, ..., d\epsilon_1$$

3 Multinomial logit.

- The multinomial logit is a model of probabilistic choice.
- In our context, with $n+1$ alternatives (including the outside "no purchase" option) the probability that a consumer ℓ chooses product i is

$$P_{\ell,i} = \frac{e^{\frac{v_i}{\mu}}}{\sum_{j=0}^{n} e^{\frac{v_j}{\mu}}}, \tag{12}$$

  where $\mu > 0$ and the outside option is alternative 0 (v_0 remains to be defined).
- μ is a scaling parameter that measures how difference in expected utilities affect choice: it can be construed as a measure of product differentiation.

The double exponential distribution

- Assume $\epsilon_{\ell,i}$ is i.i.d across consumers and products with c.d.f.

$$\bar{F}(x) = e^{-e^{-\frac{x}{\mu}-\gamma}}, \tag{13}$$

where γ is Euler's constant ($\simeq .5772$).

- ϵ has mean zero, variance $\frac{\mu^2\pi^2}{6}$ and support $(-\infty, +\infty)$.
- Density is

$$f(x) = \frac{1}{\mu} e^{-\frac{x}{\mu}-\gamma} e^{-e^{-\frac{x}{\mu}-\gamma}}. \tag{14}$$

- Relevant demand expression is (9) where the lower bound of the integral $-v_1$ is replaced by a.
- Let us define $t = e^{-\frac{\epsilon_1}{\mu}-\gamma}$ so that $dt = -\frac{1}{\mu}e^{-\frac{\epsilon_1}{\mu}-\gamma}d\epsilon_1$ and $f(\epsilon_1)d\epsilon_1 = -e^{-t}dt$ (we use this to make a change of variable in the integral.
- Also define $Y_i = e^{\frac{v_i}{\mu}}$, $i = 1, ..., n$ so we have

$$F(\epsilon_1 + v_1 - v_i) = e^{-t\frac{Y_i}{Y_1}}. \tag{15}$$

- Firm 1's demand can be written as

$$D_1(p_1, ..., p_n) = \int_0^\infty \Pi_{i=2}^n e^{-t\frac{Y_i}{Y_1}} e^{-t} dt \tag{16}$$

$$= \int_0^\infty e^{-t\frac{\sum_{i=1}^n Y_i}{Y_1}} dt \tag{17}$$

$$= \left[-\frac{Y_1}{\sum_{i=1}^n Y_i} e^{-t\frac{\sum_{i=1}^n Y_i}{Y_1}} \right]_0^{+\infty} = \frac{e^{\frac{v_1}{\mu}}}{\sum_{i=1}^n e^{\frac{v_i}{\mu}}}. \tag{18}$$

- To have an outside option and keep the multinomial logit form we need to make the associated utility random: assume that not buying any product yields:

$$u_{\ell,0} = \epsilon_{\ell,0} + y_\ell, \tag{19}$$

  where $\epsilon_{\ell,0}$ is i.i.d across consumers and with all product specific random terms $\epsilon_{\ell,i}$, $i = 1, .., n$.

- Then consumer ℓ chooses alternative $i = 0, ..., n$ with probability given by (12) with $v_0 = 0$, that is

$$P_{\ell,i} = \frac{e^{\frac{v_i}{\mu}}}{1 + \sum_{j=1}^{n} e^{\frac{v_j}{\mu}}}, \tag{20}$$

- Assume $\alpha_i = x_i\bar{\beta}$, where x_i is a vector of characteristics for product i.
- Then we can write
$$\frac{v_i}{\mu} = x_i\beta - \delta p_i, \tag{21}$$
where $\beta = \frac{1}{\mu}\bar{\beta}$ and $\delta = \frac{1}{\mu}$.
- Then individual demand can be estimated using a standard multinomial logit
- Under the assumption that $\epsilon_{\ell,i}$ is i.i.d across consumers, aggregate demand for product i is merely $D_i = LP_{\ell,i}$.

IIA property

- Take a consumer ℓ who consumes either product 1 or product 2: then the conditional probability that 1 is chosen is

$$\frac{P_{\ell,1}}{P_{\ell,1}+P_{\ell,2}} = \frac{e^{\frac{v_1}{\mu}}}{e^{\frac{v_1}{\mu}}+e^{\frac{v_2}{\mu}}}. \tag{22}$$

  It does not depend on the number of other products available or on their characteristics (v_j for $j \neq 1, 2$).
- This property is called "independence of irrelevant alternatives" (IIA)..

- This is an unappealing property: suppose the choice probabilities for 3 different cars are
  1. Renault Clio: .5
  2. Mercedes S class: .25
  3. Porsche Cayenne: .25
- IIA says that if Porsche Cayenne is taken out of the market, probabilities for Renault Clio and Mercedes S Class go up to 2/3 and 1/3 respectively (which were the conditional probabilities when the Porsche was still around).
- At the aggregate level, the market share for the Renault would increase more than that of the Mercedes.
- For a price change we have $\frac{dP_{\ell,i}}{dp_3} = \frac{1}{\mu} P_{\ell,i} P_{\ell,3}$, $i = 1, 2$, so the brand with a large share is more affected by a change in price of the third brand.

4 Oligopoly

- Assume now that each product i is produced at constant marg. cost c_i.
- Products are sold by $M \leq n$ firms.
- Assuming $\epsilon_{\ell,i}$ i.i.d over ℓ and i, $i = 0, ..., n$ demand for product i is

$$D_i = L \int_a^b \Pi_{j \neq i} F(\epsilon + v_i - v_j) f(\epsilon) d\epsilon, \tag{23}$$

with $v_0 = 0$.

- Firms choose prices simultaneously in a Nash equilibrium.
- p_i^*: equilibrium price of product i.

First order conditions: single product firms

- Assume now each firm sells one product so $M = n$ (and we call i the firm selling product i).
- Consider firm i's problem: it selects a price p_i expecting other firms to charge equilibrium prices p_j^*, $j \neq i$.
- Its demand deriv. with respect to own price p_i is

$$\frac{\partial D_i}{\partial p_i} = -L \int_a^b \sum_{j \neq i} f(\epsilon + v_i - v_j) \Pi_{k \neq i,j} F(\epsilon + v_i - v_k) f(\epsilon) d\epsilon. \tag{24}$$

4 Oligopoly
First order conditions: single product firms

- As for a single product monopolist the price FOC is

$$p_i^* - c_i = -\frac{D_i}{\partial D_i / \partial p_i}. \tag{25}$$

- Jointly estimating demand and this FOC would yield estimates for demand parameters and marginal costs.
- However actual data concern multi product firms so the price FOCs will be more complex.

First order conditions: multi product firms

- Now assume each firm m sells several products and denote K_m the set of products sold by firm m.
- Firm m's profit can be written as

$$\pi_m = \sum_{i \in K_m} (p_i - c_i) D_i. \tag{26}$$

.

- Hence, FOC for p_i is

$$p_i^* - c_i = -\frac{D_i}{\partial D_i / \partial p_i} - \sum_{j \in K_m, j \neq i} \frac{\partial D_j / \partial p_i}{\partial D_i / \partial p_i} (p_j^* - c_j). \tag{27}$$

Consumer heterogeneity in the LRUM model

- In the LRUM model, consumer heterogeneity enters in two ways:
  1. through income y_ℓ.
  2. through the random term ϵ_ℓ.
- We noted already that income plays no role because it cancels out in the choice.
  - if we introduce a vector of other individual characteristics z_ℓ that enters additively in utility as $z_\ell \gamma$ it would also be canceled out.

Consumer heterogeneity in the LRUM model

- The random term $\epsilon_{\ell,i}$ does not interact with product specific variables summarized in v_i.
- As a result, market shares are fully determined by the values of v_i independent of its composition:
  - e.g. demand for a high quality (high α_i) product with a high price could be analogous to that of a low quality product with a low price (same market share, same price derivatives).
  - observed product characteristics only matter through their contribution to α_i but it does not matter whether two products have different or similar characteristics.
- IIA for the logit is an extreme consequence of this.

- When dealing with individual choice data, individual heterogeneity and individual characteristics can be accounted for through a *mixed logit* model.
- In our econometric specification of the logit model, we could for instance assume that the price parameter is consumer specific and can be written $\delta_\ell = z_\ell \gamma + \eta_\ell$ where η_ℓ is i.i.d across consumers.
- A similar strategy could be used for the parameters in β associated with the various product characteristics.
- Then choice probabilities are obtained by taking the mean of the logit expression over the random terms.

Aggregate demand data and random coefficients

- With aggregate demand data, random coefficients can be introduced with aggregate data on the distribution of individual characteristics (e./g. distribution of income).
- Then the market share is computed as the mean over all realizations of z_ℓ of the mixed logit probabilities.