ORGANIZACIÓN INDUSTRIAL EMPÍRICA
IN7E0

Carlos Noton

Clase 4 - Jueves 9 de Agosto
Primavera - 2018

Outline

Homogenous or Differentiated Goods?

- Previous Assumption: Goods are homogenous: Bertrand paradox, Cournot competition or Collusion.
- New Assumption: Goods are differentiated in many respects.
- Differentiated products imply a non-perfect substitution by the consumer. Consumers not always choose the cheapest good, since there are other dimensions to look at.
- Consumer's imperfect substitution grants **market power** to the producer, so prices are higher than marginal cost.

Empirical IO with Differentiated Goods

Demand Estimation for Differentiated Products poses a variety of challenges. Main issues:

- Too many goods. In most markets there are several varieties of products. This implies that there are many many own- and cross-price elasticities.
- In practice, there are often too many elasticities to practicably estimate.
- Additionally, we often worry about prices being *endogenous*.
- Prices are contaminated with demand shocks (the error term), meaning that OLS estimation yields biased estimates of our elasticities of interest.

Problem 1: Too Many Goods

- Consider a typical product market like cars, cereal, or mobile phone services. While there may be a few (e.g. less than 10) firms offering such products, there may be many product varieties.
- To see why this is a problem, consider estimating a simple linear demand system:

$$Q_1 = \beta_{0,1} + \beta_{1,1} p_1 + \cdots + \beta_{J,1} p_J + \varepsilon_1$$
$$\vdots$$
$$Q_J = \beta_{0,J} + \beta_{1,J} p_1 + \cdots + \beta_{J,J} p_J + \varepsilon_J$$

- With J products, there are J^2 elasticities to estimate
- $J = 50 \Rightarrow 2,500$ elasticities to estimate.
- One would need *lots* of data to try to pin down so many parameters.

Problem 2: Price Endogeneity

- The same problem as in the homogenous good case. Prices are not randomly determined. Prices are chosen by rational producers and therefore prices should be correlated with unobservable demand shocks.
- The endogeneity problem makes the too-many-parameters problem even worse: Instead of just needing sufficient variation in prices to identify all the own- and cross-price elasticities, we now need sufficient *instruments* to identify them all.

Demand Estimation Solutions

A variety of methods have developed to solve these problems. They all rely on one or both of the following strategies:

- **Restrict the parameter space** by making all the own- and cross-price elasticities a function of a smaller number of parameters.
- **Find good instruments**: the ideal instruments for a general demand system would be cost shifters for *each* available good.

This instrument is generally unavailable, and so efforts are made to find as many plausible instruments as possible.

Different Restrictions on the Parameter Space

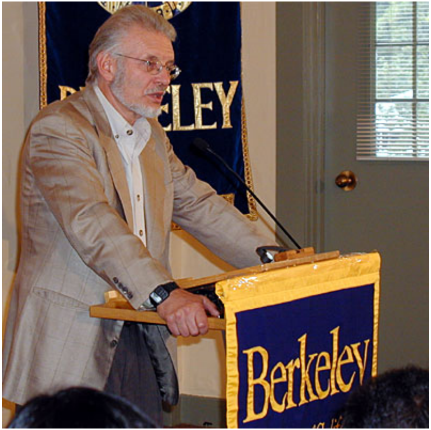
- Aggregated discrete-choice models of product differentiation.
- Assume consumers choose among discrete mutually-exclusive alternatives.
- Specify consumer utility as a function of the observed and unobserved characteristics of the product. (Lancaster 1966)
- Aggregate demand over consumers to get product level market shares.

There are three main generations of models:

- Logit Model (70's)
- GEV Model (80's)
- Discrete-Choice Random Coefficients Model (RCM) (since 95)

Logits

Simple things first: Logit developed by Daniel L McFadden.



Logit Model

Let the conditional indirect utility function for consumer i for product j in market (or time period) n be

$$u_{ijn} = \alpha(y_i - p_{jn}) + x_{jn}\beta + \xi_{jn} + \varepsilon_{ijn}$$

where

- p_{jn}: price of j in n
- y_i: income of individual i
- x_{jn}: observed (to econometrician) characteristics of j in n
- ξ_{jn}: unobserved characteristic of or demand shock for j in n
- α, β: common parameters, ie, identical for all consumers.
- The constant αy_i does not play a role when comparing heterogenous goods.
- ε_{ijn}: idiosyncratic taste consumer i has for j in n, which are *iid*

Logit Model Notes

$$u_{ijn} = \alpha(y_i - p_{jn}) + x_{jn}\beta + \xi_{jn} + \varepsilon_{ijn}$$

Note:

- We are assuming the utility from product j depends on the characteristics it offers, x_{jn}, since they are differentiated products demand.
- We could have linear income in our indirect utility function. (They drop in utility comparisons).
- If individual level data of income is available, we could have *income effects*.

Logit Model Notes

$$u_{ijn} = \alpha(y_i - p_{jn}) + x_{jn}\beta + \xi_{jn} + \varepsilon_{ijn}$$

Note:

- Notice there are two random shocks: ξ_{jn} is an *unobserved* (to the econometrician) product characteristic (or demand shock); ε_{ijn} is the individual level shock.
- ξ_{jn} is the source of our endogeneity problem.
- Why the endogeneity?
- Because ξ_{jn} is likely to be observed by both consumers and firms. Firms set prices as a function of ξ_{jn}!
- Products that have higher ξ_{jn} are more appealing to consumers and firms take advantage of that charging higher prices (examples: adv campaign, hidden quality, fashion, etc)

Deriving the Logit Model

Define:

$$\delta_{jn} = -\alpha p_{jn} + x_{jn}\beta + \xi_{jn}$$

δ_{jn} is the "mean utility" for j in n (notice that there is no i here, same for all individuals).
What is the probability that consumer i choose product j? $\Leftrightarrow$ The probability good j yields the highest utility in this **discrete** choice, hence:

$$\begin{aligned}
\mathbb{P}(i \text{ buy } j) &= \mathbb{P}(u_{ijn} > u_{ikn}, \forall k \neq j) \\
&= \mathbb{P}(u_{ijn} > u_{i1n}, .., u_{ijn} > u_{ij-1n}, u_{ijn} > u_{ij+1n}, .., u_{ijn} > u_{iJn} \\
&= \mathbb{P}(\delta_{jn} + \varepsilon_{ijn} > \delta_{kn} + \varepsilon_{ikn}, \forall k \neq j) \\
&= \mathbb{P}(\delta_{jn} - \delta_{kn} > \varepsilon_{ikn} - \varepsilon_{ijn}, \forall k \neq j)
\end{aligned}$$

Deriving the Logit Model

Denote $\widetilde{\varepsilon}_{inj} = (\varepsilon_{in1} - \varepsilon_{inj}, .., \varepsilon_{inj-1} - \varepsilon_{inj}, \varepsilon_{inj+1} - \varepsilon_{inj}, .., \varepsilon_{inJ} - \varepsilon_{inj}$ is the $J-1$ vector of the error differences between product j and all the other products for consumer i.

$$\begin{aligned} \mathbb{P}(i \text{ buy } j) &= \mathbb{P}(\delta_{jn} - \delta_{kn} > \varepsilon_{ikn} - \varepsilon_{ijn}, \forall k \neq j) \\ &= \int 1\left(\widetilde{\varepsilon}_{inj} < \delta_{jn} - \delta_{kn}, \forall k \neq j\right) g(\widetilde{\varepsilon}) d\widetilde{\varepsilon} \end{aligned}$$

Function $1\left(a\right)$ is equal to one if statement a is true and zero otherwise (Indicator Function).

Now what?!! How to deal with this awful integral!

Deriving the Logit Model

In the 70's, McFadden developed a clever way how to deal with these unfriendly integrals.

If $\varepsilon_{ijn} \sim$ Type I Extreme Value, the difference $\widetilde{\varepsilon}_{inj}$ distributed with a logistic distribution. Moreover:

$$\begin{aligned}\mathbb{P}(i \text{ buy } j) &= \frac{e^{\delta_{jn}}}{\sum_{k=0}^{J} e^{\delta_{kn}}} \\ &= \frac{e^{-\alpha p_{jn} + x_{jn}\beta + \xi_{jn}}}{\sum_{k=0}^{J} e^{-\alpha p_{kn} + x_{kn}\beta + \xi_{kn}}}\end{aligned}$$

Note:

- This expression is at individual level.
- We still have a bunch of unobservable ξ_{jn} that are correlated with price.

Some standard normalizations

When dealing with discrete-choice models, there are two standard normalizations that one must make:

- Utilities and affine transformations. if u works, $a \times u + b$ with $a > 0$ also works.
- Setting the utility to one of the goods to zero. All that matters is *differences* in utility.
- Thus we set one utility to zero and measure all other utility *relative* to this baseline.
- We denote the "outside good," by $j = 0, u_{i0n} = \varepsilon_{i0n}$ (Zero characteristics, zero price, "not buying any good" $\Rightarrow \delta_{0n} = 0$
- Also, we "set the scale of utility" by normalizing the variance of ε.

Estimation of Logit Models

We observe market shares. We compute probabilities. How to relate them?

The market share is the proportion of the population that purchased product j. It must be equal to the probability that the population choose that alternative. **We can identify our parameters by matching probabilities with market shares!**

$$\mathbb{P}(population \text{ choose } j) = s_{jn}$$

Hence, we need to aggregate every consumer to have at population level probabilities.
Adding over the population that prefers j, you can derive the theoretical market shares:

In the logit model, there is no heterogeneity among consumers ($s_{jn} = s_{ijn}$).

Logit Model and Market Shares

In the logit model there is no heterogeneity, hence the market share is the probability of the representative consumer:

$$
\begin{aligned}
s_{jn} &= \mathbb{P}(i \text{ buy } j) \\
&= \frac{e^{\delta_{jn}}}{1+\sum_{k=1}^{J} e^{\delta_{kn}}} \\
&= \frac{e^{-\alpha p_{jn}+x_{jn}\beta+\xi_{jn}}}{1+\sum_{k=1}^{J} e^{-\alpha p_{kn}+x_{kn}\beta+\xi_{kn}}}
\end{aligned}
$$

Recall that the outside good has $\delta_0 = 0$, hence $e^{\delta_0} = e^0 = 1$ (this is the first term in the denominator!). Therefore, the fraction of the outside good is $s_{0n} = \frac{e^{\delta_0}}{1+\sum_{k=1}^{J} e^{\delta_{kn}}} = \frac{1}{1+\sum_{k=1}^{J} e^{\delta_{kn}}}$.

Logit Model and Market Shares

Recall that we want to estimate the utility parameters:

$$
\begin{aligned}
\log\left(\frac{s_{jn}}{s_{0n}}\right) &= \log\left(\frac{e^{\delta_{jn}}}{1+\sum_{k=1}^{J} e^{\delta_{kn}}} \div \frac{1}{1+\sum_{k=1}^{J} e^{\delta_{kn}}}\right) \\
&= \log\left(e^{\delta_{jn}}\right) \\
&= \delta_{jn} \\
&= -\alpha p_{jn} + x_{jn}\beta + \xi_{jn}
\end{aligned}
$$

Having ξ_{jn} in a linear relationship is a great thing!! Otherwise it is complicated to use IV approach...

Endogeneity Issue

We have the following linear regression model:

$$\log\left(\frac{s_{jn}}{s_{0n}}\right) \;=\; -\alpha p_{jn} + x_{jn}\beta + \xi_{jn}$$

Can I use OLS?

No. Recall the endogeneity problem!

Still need instruments (say Z_{jn}). We need predicted prices $\widehat{p}_{jn} = f(Z_{jn})$ that are not contaminated with ξ_{jn}.

We will discuss suitable instruments later on...

Logit Elasticities

Finally, we can estimate elasticities in a demand for differentiated products:

- **Own-price derivatives / elasticities**:

$$\frac{\partial s_{jn}(p_n)}{\partial p_{jn}} = \frac{\partial s_{jn}(p_n)}{\partial \delta_{jn}} \frac{\partial \delta_{jn}}{\partial p_{jn}} = s_{jn}(1 - s_{jn})(-\alpha)$$

$$\Rightarrow \epsilon_{jj} = \frac{\partial s_{jn}(p_n)}{\partial p_{jn}} \frac{p_{jn}}{s_{jn}} = -\alpha p_{jn}(1 - s_{jn})$$

- **Cross-price derivatives / elasticities**:

$$\frac{\partial s_{jn}(p_n)}{\partial p_{kn}} = \frac{\partial s_{jn}(p_n)}{\partial \delta_{kn}} \frac{\partial \delta_{kn}}{\partial p_{kn}} = s_{jn} s_{kn} \alpha$$

$$\Rightarrow \epsilon_{jk} = \frac{\partial s_{jn}(p_n)}{\partial p_{kn}} \frac{p_{kn}}{s_{jn}} = \alpha p_{kn} s_{kn}$$

Is the $i.i.d.$ assumption realistic?

- The assumption TIEV implies independent and identically distributed (i.i.d.).
- *iid* shocks across consumers and across markets seem plausible.
- However, *iid* shocks across products are quite wrong assumptions, since similar goods should have similar taste shocks (correlation between shocks of similar goods)

Independence of Irrelevant Alternatives (IIA)

- $\epsilon_{jk} = \alpha p_{kn} s_{kn}$
- Note that the cross-price elasticity of good j with respect to good k *is independent of* j!
- **Regardless of the characteristics of the goods**, the elasticity between two products depends on their market shares only.
- The unattractive property of the Logit model is called Independence of Irrelevant Alternatives (IIA) property:
  - It says that the ratio of market shares of two products depends *only* on the relative utility of *those* two products:

$$\frac{s_{jn}}{s_{mn}} = \frac{e^{\delta_{jn}}}{1+\sum_k e^{\delta_{kn}}} \frac{1+\sum_k e^{\delta_{kn}}}{e^{\delta_{mn}}} = \frac{e^{\delta_{jn}}}{e^{\delta_{mn}}}$$

- Suppose two equally popular -although very different in terms of characteristics- products with fifty percent of the market each $s_A = s_B = 0.5$.
- Consistently, the logit model will rationalize this fact with identical size of deltas: $\delta_A = \delta_B$ and *iid* shocks.
- **If a new good, C, which is identical to A, enters the market**, we should expect that market share of B remains constant in fifty percent. Market shares of A and C should add up to fifty percent.
- However, the model will keep a constant ratio of market shares between A and B.
- *Forcing* the model to predict $s_A = s_B = s_C = 0.33$, and $\delta_A = \delta_B = \delta_C$.
- Clearly, $\delta_A = \delta_C$, but we should expect that taste shocks are correlated between A and C.

Generalizing Discrete Choice Models

The general equation for discrete choice models are of the form:

$$\begin{aligned} s_{jn} &= \int_{A_j} \mathbb{P}(i \text{ buy } j) f(i) di \\ &= \int_{A_j} \left[\int 1 \left(\widetilde{\varepsilon}_{inj} < \delta_{jn} - \delta_{kn}, \forall k \neq j \right) g(\widetilde{\varepsilon}) d\widetilde{\varepsilon} \right] f(i) di \end{aligned}$$

- We saw particular functional forms for Logit and Nested Logit.
- There are many options for $g(\cdot)$: probits (where g is normal-multivariate), generalized extreme value functions (GEV), Multinomial, and a large list of etceteras. (see Train's book for a great survey).

Generalizing Discrete Choice Models

$$s_{jn} = \int_{A_j} \left[\int 1 \left(\widetilde{\varepsilon}_{inj} < \delta_{jn} - \delta_{kn}, \forall k \neq j \right) g(\widetilde{\varepsilon}) d\widetilde{\varepsilon} \right] f(i) di$$

- If heterogeneity between consumers play a role, then $f(\cdot)$ is important and $\mathbb{P}$ is different across consumers: Random Coefficient Models or Mixed Logits to estimate (α_i, β_i)
- The functional form for δ_{jn} also matters, so far linear but it could be extended as well.
- If you have individual level data, then you can have individual level probabilities, and the usual way to estimate is with Maximum Log-likelihood estimation (MLE).

MLE with Individual Level Data

Suppose you have individual decisions y_i and can compute the individual probabilities s_{ij}:

$s_{ij} = Pr[y_i = j] \;\; \forall j = 1,..,J$ (goods) and $i = 1,..,N$ (individuals) .

Then the estimation will maximize $\mathfrak{L}$

$$L = \prod_{i=1}^{N}\prod_{j=1}^{J} s_{ij}^{y_{ij}} \Rightarrow \mathfrak{L} = \ln L = \sum_{i=1}^{N}\sum_{j=1}^{J} y_{ij} ln(s_{ij})$$

where the dummy y_{ij} is one if individual i chose product j and zero otherwise.

We will explore non-parametric approaches to Mixed Logits with individual level data.

Instruments for Demand of Differentiated Products

Remember the simplest logit regression.

$$\log\left(\frac{s_{jn}}{s_{0n}}\right) \;=\; -\alpha p_{jn} + x_{jn}\beta + \xi_{jn}$$

- Suppose characteristics, x_{jn}, are exogenous (predetermined and insensitive to ξ_{jn} (demand shock or unobservable characteristic).
- Price instead is endogenous: expected correlation? $cov(p_{jn}\xi_{jn}) > 0$.
- Looking for suitable instruments Z_{jn} such that: i) $\mathbb{E}(Z_{jn}p_{jn}) \neq 0$; and ii) $\mathbb{E}(Z_{jn}\xi_{jn}) \neq 0$
- The literature has suggested 3 candidates: i) prices in other markets; ii) cost shifters; and iii) Characteristics of other products in the same firm and competing firms.

i) prices in other markets

Let $Z_{jn} = p_{jm}$ with $m \neq n$

- Is $\mathbb{E}(Z_{jn} p_{jn}) \neq 0$?
- Is $\mathbb{E}(Z_{jn} \xi_{jn}) \neq 0$?
- Under which circumstances these instruments are wrong?

Please see long and nasty discussion between Hausman and Bresnahan.

ii) cost shifters

Let $Z_{jn} = c_{jm}$

- Is $\mathbb{E}(Z_{jn} p_{jn}) \neq 0$?
- Is $\mathbb{E}(Z_{jn} \xi_{jn}) \neq 0$?
- Under which circumstances these instruments are wrong?

iii) Characteristics of other products in the same firm and competing firms.

Let $Z_{jn} = \sum_{k \neq j} x_{km}$ for the same firm and $\sum_{k \neq j} x_{km}$ for competitors.

- Is $\mathbb{E}(Z_{jn} p_{jn}) \neq 0$?
- Is $\mathbb{E}(Z_{jn} \xi_{jn}) \neq 0$?
- Under which circumstances these instruments are wrong?

Conclusions

- To deal with demand of differentiated products, we introduce discrete choice models.
- These models allows to reduce the dimensionality of the problem from the space of prices to the space of characteristics.
- The problem leads to a multidimensional integral that represents the probability of purchase.
- The optimization problem is to match predicted and actual market shares.
- We discussed the suggested instruments for price in the context of differentiated products.