1 Demand Estimation

1. Motivation.

2. Estimation.

3. Consistent Estimation

4. Limitations of Logit.

5. BLP

2 Motivation

- We begin our study of differentiated product markets by describing the method of BLP (1995) for demand estimation in differentiated product markets.

- We will also discuss some limitations of this method and some possible extensions.

- BLP is a method for estimating demand in differentiated product markets using aggregate data.

- The method allows for endogenous prices and random coefficients.

- The method also allows for consistent estimation of the model parameters even if there is imperfect competition.

- Next week, we will discuss flexible methods using microdata.

3 A simple example

- To motivate the framework, consider the following simple example based on Berry (RAND, 1994).

- There are $i = 1, ..., I = \infty$ agents in $t = 1, ..., T$ markets.

- Remark: It is important that there are an infinite number of agents.

- Each agent makes a choice between $j = 1, ..., J$ mutually exclusive alternatives.

- $x_{j,t} = (x_{jt,1}, ..., x_{jt,K})'$ is a $K \times 1$vector of characteristics for product j.

- Product characteristics/choice sets may evolve over markets.

- Let $p_{j,t}$ denote the price of j at time t.

- $\xi_{j,t} = \xi_j + \xi_t + \Delta\xi_{j,t}$ denote an unobserved characteristic/demand shock/measurement error in price.

- ξ_j is a permanent component for j, ξ_t is a common shock and $\Delta\xi_{j,t}$ is a product/time specific shock for j.

- Specify the random utility as:

$$u_{ijt} = x'_{j,t}\beta - \alpha p_{j,t} + \xi_{j,t} + \varepsilon_{ij}$$

- Assume that the error term is iid with density:

$$\exp(-\varepsilon + \gamma) \exp\left[- \exp\left(-\varepsilon + \gamma\right)\right]$$

- Then the market share for j at time t is:

$$s_{jt}(x, \beta, \alpha, \xi) = \frac{\text{exp}(x'_{j,t}\beta - \alpha p_{j,t} + \xi_{j,t})}{\sum_{j'=1}^{J} \text{exp}(x'_{j',t}\beta - \alpha p_{j',t} + \xi_{j',t})}$$

- Berry assumes that we are working with aggregate data and that, at the true parameter values, $s_{jt}(x, \beta, \alpha, \xi) = S_{jt}$ where S_{jt} denotes the "true" market share.

- This is why $I = \infty$ is important

- This differs from the standard conditional logit model in two ways.

- First, we have unobserved heterogeneity/demand shock, $\xi_{j,t}$.

- Why $\xi_{j,t}$?

1. Observe list of product attributes is incomplete. This goes back to hedonic regressions.

2. Measurement error in prices. Typically price data is an average.

3. Without $\xi_{j,t}$, shares should not vary holding $x'_{j,t}\, p_{j,t}$ fixed. This is likely to be violated in some data sets.

- Second, we are working with aggregate data instead of individual choices, as in the standard conditional logit.

- Thus, the data set needs to contain market shares.

- What if market shares are measured with error?

3.1 Estimation.

- Berry notes that the following transformation can be made:

$$\log(s_{jt}(x,\beta,\alpha,\xi)) = e_t + x'_{j,t}\beta - \alpha p_{j,t} + \xi_{j,t}$$
$$e_t = -\log(\sum_{j'=1}^{J} \exp(x'_{j',t}\beta - \alpha p_{j',t} + \xi_{j',t}))$$

- Next we assume a "law of large numbers" so that $S_{jt} = s_{jt}(x,\beta,\alpha,\xi)$ at the true parameters.

- If we normalize the utility of the outside good to zero, this implies that:

$$\begin{aligned} s_{0t}(x,\beta,\alpha,\xi) &= \frac{\exp(0)}{\sum_{j'=1}^{J} \exp(x'_{j',t}\beta - \alpha p_{j',t} + \xi_{j',t})} \\ \log s_{0t}(x,\beta,\alpha,\xi) &= 0 - e_t \end{aligned}$$

- This implies that:

$$\log(S_{jt}) - \log(S_{ot}) = x'_{j,t}\beta - \alpha p_{j,t} + \xi_{j,t}$$

- where S_{ot} is the share of the outside good.
- Berry noted that an obvious way to estimate this model is by regression.
- The dependent variable is $\log(S_{jt}) - \log(S_{ot})$
- The independent variables are $[x'_{j,t}, p_{j,t}]$
- The error term is $\xi_{j,t}$.
- However, in general we would expect $cov(p_{j,t}, \xi_{j,t}) \neq 0$.

- In the presence of a demand shock, oligopoly models suggest that firms should raise prices.

- Thus, ols estimates of β and α will be biased.

3.2 Consistent Estimation.

3.2.1 Fixed Effects

- A first approach to consistent estimation would be to estimate the following fixed effects model:

$$\log(S_{jt}) - \log(S_{ot}) = x'_{j,t}\beta - \alpha p_{j,t} + \xi_j + \xi_t + \Delta\xi_{jt}$$

- Where ξ_j is a brand fixed effect, ξ_t is a category market/time shock

- The identifying assumption is $E[\Delta\xi_{jt}|x'_{j,t}, p_{j,t}] = 0$

- This is clearly more appealing that $E[\xi_{jt}|x'_{j,t}, p_{j,t}] = 0$

- However, there are a couple of limitations.

- First, there may be colinearity between ξ_j and $x_{j,t}$ if some characteristics for product j are time invariant.

- Thus, a brand fixed effect does not allow us to learn about the valuation of individual product characteristics.

- Also, it presumes that $cov(p_{jt}, \Delta\xi_{jt}) = 0$

- This assumes that in a given time period, product level price variation is exogenous.

- Remark: This type of assumption is commonly made in marketing.

3.2.2 BLP Instruments

- A second approach to identification is to find a set of instruments.

- That is, we need to find a variable z_{jt} such that $E[\xi_{jt}|z_{jt}] = 0$, cov$(z_{jt}, [x_{jt}, p_{jt}])\neq 0$ (i.e. satisfies standard rank conditions for IV).

- One obvious instrument is a supply shifter (e.g. change in costs).

- Problem, there are too few instruments and they may be weak.

- Weak instruments- standard errors incorrect, bias large.

- BLP and Berry(1994) suggest measures of isolation in product space.

- e.g. $z_{jtk} = \sum_{j' \neq j} x_{j'tk}$

- How much does product j contribute to the (unweighted) average of characteristic k.

- This instrument is usually available and it tends to be highly correlated with price.

- Models of oligopoly suggest the more isolated you are in product space, the more likely you are to have a higher margin.

- Thus, prices will be correlated with z_{jtk}.

- Critiques of this instrument-

1. Little variation over time.

2. Assumes $cov(\xi_{jt}, x_{jtk}) = 0$.

- This assume that omitted product attributes are uncorrelated wth observed attributed.

- This seems hard to believe since the observed attributes are correlated with each other.

- This is a classic problem in demand estimation.

- In hedonic, researchers have long worried about the consistency of:

$$p_{jt} = x'_{jt}\beta + \xi_{jt}$$

- For example, in a home price regression, the observed attributes are likely to be correlated with the unobserved attributes.

- Ackerberg, however, notes that if $cov(z_{jtk}, x_{jtk}) = 0$ for all k, it is possible to consistently estimate price elasticities for this model (even if other parameter estimates are biased).

- This condition is testable!

- Many questions can be answered with price elasticities- e.g. measurement of market power.

- As with the the fixed effects case above, it seems more appealing to assume:

$$E[\Delta\xi_{jt}|z_{jt}] = 0$$

- This is possible if we include brand/time fixed effects.

- Remark: Price endogeneity is being accounted for using only demand side information.

3.3 Hausmann Instruments

- Hausmann proposes using prices in other markets as instruments.

- E.g. use prices in Iowa, Wisconsin and the Dakotas as instruments for price endogeneity in Minneapolis.

- The idea behind these instruments is that they pick up common cost shocks.

- However, if they pick up common demand shocks, they are invalid.

- In general, both the BLP and Hausmann instruments have the advantage of at least being available!

4 Limitations of the Logit

- Some Limitations of the Logit

- While the logit model is computationally convenient, it imposes some unpleasant restrictions on the data.

- It is still widely used since there are few other computationally convenient estimators.

1. Implausible substitution patters.

- In the logit model exhibits the independence of irrelevant alternatives (IIA).

- That is, the ratio of the probability of two choices does not change depending on the set of choices that are available.

$$\frac{\text{Pr(i chooses j)}}{\text{Pr(i chooses j}')} = \text{constant}$$

for all j and $j\prime$ regardless of the set of alternatives that are available.

- A famous example is the red bus/blue bus problem

- Suppose that we are studying the mode of transportation choice.

- Choice set is take the (red) bus to work or to drive.

- Suppose that these choices are equal in probability.

- Now suppose that the bus company introduces blue buses in addition to red buses.

- Suppose that consumers are indifferent about the color of their bus and that the probability of the red bus and blue bus is equal.

- IIA implies that prob(red bus)$=$ prob(blue bus) $=$prob(drive) $=1/3$.

- A more "intuitive" answer would be rob(red bus)= prob(blue bus)= 1/4 and prob(drive)=1/2.

- This example shows that IIA can give wierd substitution patterns.

- This can also show up in terms of price elasticities.

- Suppose that we are modeling consumer demand for a differentiated product.

- Suppose that the latent utilities are:

$$y_{nj} = x'_{nj}\beta - \alpha p_j + \varepsilon_{nj}$$

- where p_j is price.

- Calculating the own and cross price elasticites.

$$\begin{aligned}\eta_{jk} &= \frac{\partial \Pr(\text{i chooses j})}{\partial p_k}\frac{p_k}{\Pr(\text{i chooses j})} \\ &= \begin{cases}-\alpha p_j(1-s_j) \text{ if } j=k \\ -\alpha p_k s_k\end{cases}\end{aligned}$$

- Since in most cases there are many products so that the market shares are typically small, $(1-s_j)$ is approximately equal to price.

- This implies that the lower the price the lower the elasticity.

- This implies that markups should be higher in cheap products.

- This is clearly not appropriate in many industries.

- A second limitation is that cross price elasticities are determined by $\alpha p_k s_k$.

- Suppose that Lucky Charms and Grape Nuts are similarly priced and have a similar market share.

- An implication of this formula is that both of these will have the same cross price elasticity with CoCo Puffs.

- This is clearly a priori implausible, yet it is an assumption that we have imposed through the functional form.

3. Treatment of Heterogeneity.

- In the logit model, consumers are only heterogenous because of ε_{ij}.

- ε_{ij} can be thought of as adding additional product characteristics into the model for each j and an iid random preference shock for that characteristic.

- Caplin and Nalebuff argue that this generates too much "taste for variety".

- Applied studies of welfare, such as Petrin (JPE 2002, Quantifying the Benefits of New Products: The Case of the Minivan), argue that to much of the utility comes from implausbly large draws of the ε_{ij}.

- Leads to pathological implications (e.g. markups in Bertrand may not converge to zero as market becomes thick).

- See Anderson, DePalma and Thisse.

5 BLP-Random Coefficients Logit.

- In Berry (1994) and BLP (1995), consumer preferences can be written as:

$$u(x_j, \xi_j, p_j, v_i; \theta_d)$$

where:

- $x_j = (x_{j,1}, ..., x_{j,K})$ is a vector of K characteristics of product j that are observed by both the economist and the consumer.

- ξ_j is a characteristic of product j observed by the consumer but not by the economist.

- p_j is the price of good j

- v_i vector of taste parameters for consumer i
- θ_d vector of demand parameters.
- One commonly used specification is the logit model with random (normal) coefficients:

$$u_{ij} = x_j \beta_i - \alpha p_j + \xi_j + \varepsilon_{ij}$$

- The K random coefficients are:

$$\begin{aligned} \beta_{i,k} &= \beta_k + \sigma_k \eta_{i,k} \\ \eta_{i,k} &\sim N(0,1),\ iid \end{aligned}$$

- Consumer i will purchase good j if and only if it is utility maximizing, just as in the previous lecture.

- Question: How do we interpret the parameters of this model?

- It is useful to decompose utility into two parts, the first is a "mean" level of utility and the second is a heteroskedastic error terms that captures the effect of random tastes parameters:

$$\begin{aligned} \upsilon_{ij} &= \left[\sum_k x_{jk} \sigma_k \eta_{i,k} \right] + \varepsilon_{ij} \\ \delta_j &= x_j \beta - \alpha p_j + \xi_j \end{aligned}$$

- We can now write utility of person i for product j as:

$$u_{ij} = \delta_j + \upsilon_{ij}$$

- Next, we will write the market shares for aggregate demand in a particularly convenient fashion. First define the set of "error terms" that make product j utility maximizing given the J dimensional vector $\delta = (\delta_j)$

$$A_j(\delta) = \left\{ \upsilon_i = (\upsilon_{ij}) | \delta_j + \upsilon_{ij} \geq \delta_{j'} + \upsilon_{ij'} \text{ for all } j' \neq j \right\}$$

- The market share of product j can then be written as (assuming a law of large numbers):

$$s_j(\delta(x, p, \xi), x, \theta) = \int_{A_j(\delta)} f(\upsilon) d\upsilon$$

- In this case, the parameter θ is β, α and σ.

- Given θ and the demand for product j actually observed in the data, $\widetilde{s}_j$ it must be the case that:

$$\widetilde{s}_j = s_j(\delta(x, p, \xi), x, \theta)$$

- Given θ, this can be expressed as a system of J equations in J unknowns (the ξ_j).

- To estimate, we find a set of instruments for the ξ_j.

- We must find a set of instruments correlated with the endogenous variable p_j, but uncorrelated with the residual ξ_j.

Commonly used instruments:

1. The product characteristics.

2. Prices of products in other markets (interpret ξ_j as a demand shifter).

3. Measures of isolation in product space ($\sum_{j' \neq j} x_{j',k}$)

4. Cost shifters.

- Question: Are these really valid instruments?

- Typically we think of product characteristics as a choice variable.

- Suppose that a firm chose product characteristics optimally.

- Then the unobserved characteristics (to the econometrician) of a product would be independent of the observed characteritics only under strong separability assumptions about cost and demand.

- The model written down probably violates the separability assumptions on demand.

- A number of empirical case studies have been done. They find that BLP style estimators typically find more elastic demand curves.

6 Firm Behavior.

- In the model above, we abstracted from the behavior of the firm.

- Suppose that firms engage in Bertrand price competition.

- Let firm f produce some set of products P_f.

- Then to profit maximization problem for firm f is to choose prices p_j for $j \in P_f$ that maximize expected profit holding the prices of the other firms fixed:

$$\pi_f = \sum_{j \in P_f} (p_j - mc_j) M s_j(x, p, \xi, \theta)$$

- Suppose that we know the function s_j, then the first order conditions for all of the products are a system of J equations in J unknowns where the unknowns are the latent cost parameters mc_j.

- Note that if we recover the marginal cost parameters by assuming Bertrand price competition and that the first order conditions hold, we could do policy experiments.

- For instance, some have used this approproach to simulate the effects of a merger.

- BLP (1995) proceeds in a similar fashion to Berry, except that it models the supply side as well by assuming that firms are Bertrand price competitors.

- We then need to find instruments for a set of unobserved supply shifters.

- BLP propose the use of product characteristics.

7 Computation.

- In this section, I shall outline some of the key steps needed to actually compute Berry (1994).

- A key step in many programming projects is to do a fake data experiment.

- Simulate the model using fixed parameter values.

- Pretend you don't know the parameter values and estimate.

- This tests the code and sometimes shows you limitations of the models.

- One of the best ways to really learn the econometrics in a paper is to do a fake data experiment.

- We shall consider as an example the random coefficinet logit model.

There are basically 4 things we need to do in order to compute the value of the objective function in order to do GMM.

1. For a given value of σ and δ, compute the vector of market shares.

2. For a given value of σ, find the vector δ that equates the observed market shares and those predicted by the model using the contraction mapping.

3. Given δ and β, α compute the value of ξ

4. Search for the value of ξ that mimizes the objective function.

- We shall consider these one at a time.

7.1 Computing Market Shares.

- In the random coefficient logit model, we can compute the market shares, given δ as follows:

$$s_j(\delta, \sigma) = \int \frac{\exp(\delta_j + \sum_k x_{j,k}\eta_{i,k}\sigma_k)}{1 + \sum_{j'} \exp(\delta_{j'} + \sum_k x_{j',k}\eta_{i,k}\sigma_k)} df(\eta_i)$$

- In practice, the integral above is computed using simulation.

- Make a set of S simulation draws and keep them fixed for the whole problem.

- Sometimes importance sampling is useful in order to improve the speed/accuracy of the integration.

- See Judd for an overview of numerical integration.

- We can compute confidence intervals using standard methods to see whether the simulated market shares are well estimated.

7.2 The contraction mapping.

- Next, we wish to find the δ that matches the observed market shares given σ.

- In Berry and BLP they demonstrate that the following is a contraction:

$$\delta_j^{(n+1)} = \delta_j^{(n)} + \ln(\widetilde{s}_j) - \ln(s_j(\delta, \sigma))$$

- Therefore, given that we can compute market shares, we can use the formula above to find the value of δ by making an initial guess at δ and then evaluting the equation above until convergence is (approximately) achieved.

- A mapping T that maps $S \rightarrow S$ is a contraction with modulus β if for all x, y $d(T \circ x, T \circ y) \leq \beta d(x, y)$.

- A contraction mapping has a unique fixed point.

- Let v_o be an initial guess about the fixed point v. Let $T^n(v_o)$ denote applying the mapping n times, as in the previous equation.

- This converges to the fixed point at an exponential rate.

- Point: Market shares can be inverted very quickly in a fairly simple manner!

- Contraction mappings are used all the time in economics, particularly in modern Macro.

- See Stokey and Lucas, chapters 4 and 5 for proofs.

8 Computing the value of ξ

- The next set is simple. Just let:

$$\xi_j = \delta_j - (x_j\beta - \alpha p_j)$$

where δ_j is computed using the contraction mapping.

8.1 Computing the value of the objective function.

- Let Z be the set of instruments.

- The objective function is formulated as in all GMM problems assuming $E(\xi|Z) = 0$.

- The econometrician then chooses β, α, and σ in order to minimize the objective function.

- Standard mathematical programs (MATLAB, GAUSS, IMSL,NAG) contain software for optimization problems.

- One standard way to proceed is to do a rough global search first and then use a derivative based method second once you have a very rough sense of the overall shape of the objective function.

- Multiple starting points commonly used in order to search for multiple local solutions to minimization problem..

- See Judd for an overview of numerical minimization.

- Doing a "fake data experiment" is a good way to learn how well the estimator works.

- Fix true parameters, simulate the model. Then see if your computations allow you to get back the correct answer.

9 Benefits of the Minivan

1. Motivation.

2. Demand Model.

3. Data/Estimation.

4. Results

10 Motivation

- In this paper, Petrin attempts to measure the benefits from a new good- the minivan.

- Theory has ambiguous predictions about new products.

- Excessive product introduction in order to gain market power.

- New products may alternatively better match consumer needs.

- Who benefits most from new products- firms from market power/ consumers from improved products?

- The minivan was one of the most important new products in the auto industry in the 1980's.

- Interesting and important innovation (autos are about 12 percent of the CPI).

- Modifies BLP approach to allow for a combination of demographic data (from CEX) and market share data.

- This paper is also important as a critique of the logit model.

- It is truly shocking how bad the logit estimates are in this paper.

- A more advanced econometric approach, usign recent structural methods gives much more plausible results.

- This is a very well written paper (helps to explain why it is well cited).

- Clearly explains problem, techniques and results to a general applied audience.

11 Minivan Innovation

- Introduced in 1984 by Chrsler- Dodge Caravan.
- Instant success with sales of 170K in 1st year.
- GM-Astro/Safari and Ford-Aerostar in 85/86.
- Clones weren't as good becaue they were buit on truck platforms.
- Chrysler dominated this market for many years-14 years later 44% share.
- Minivan canabalized station wagon sales.
- 950K in 1984 to 300K 7 years later.

12 Demand Model

- $j-$product and $i-$consumer, parameters, θ

$$u_{ij} = \alpha_i \ln(y_i - p_j) + X_j\beta + \sum_k \gamma_{ik} x_{jk} + \xi_j + \varepsilon_{ij}$$

- p_j, price,X_j characteristics,ξ_j omitted attributes

- α_i is a spline in income to allow for flexibility in price elasticities

- $\gamma_{ik} = \gamma_k v_{ik}$ (where v_{ik} is random coefficient)

- $\gamma_{i,mi} = \gamma_k v_{ik} \ln(fs_i)-$ fs_i is family size and "mi" denote minvan

- $\gamma_{i,sw} = \gamma_k v_{ik}\ln(fs_i)$– "$sw$" denotes station wagon

- Thus, Amil allows larger families (presumably with kids) to prefer minivans

- It may be hard to learn much about random coefficients from aggregate data.

- Micro theory suggests that to learn about an individual consumer's preferences, we have to see how the consumer's choice vary as a function of income.

- Little information about individual level substitution in aggregate data.

- Amil uses consumer expenditure survey (CEX)

- Note that the raw CEX survey is not completed by many households
- Measurement error?
- This is a widely used data set from the BLS (expenditure shares for CPI)
- From BLS, it is possible to learn the covariance between demographics and vehicle charactersitics.
- This micro data gives information about how substitution patterns vary with demographic groups.
- The supply side is assumed to be multi product Bertrand price competition
- This assumption is used in counterfactuals (uniqueness???)

13 Data

- 916 vehicles marketed from 1981 to 1993
- Automotive News Market Data Book and Ward's Automotive Yearbook
- Characteristics-acceleration, vehicle dimensions, drive type, fuel efficiency, luxury (air conditioning standard)
- Prices and quantities
- CEX auto supplement, link demographics to new vehicle purchases
- Note that the number of observations on minivan and station wagon purchases are thin

14 Estimation

- The estimator is essentially BLP with additional moments on purchases conditional on demographics

$$
\begin{aligned}
&E[\text{i purchases new vehicle}|\text{income}]\\
&E[fs_i|\text{i purchases minivan}]\\
&E[fs_i|\text{i purchases station wagon}]\\
&etc...
\end{aligned}
$$

- As in our discussion in class, Amil uses the contraction mapping to numerically solve for $\delta_j(\theta)$

$$s_j = s_j(\delta_j(\theta), \theta) \ j = 1, ..., J$$

- $\xi_j(\theta)$ is computed as a residual (also incorporates $w_j(\theta)-$ shock to marginal cost)

- The moment conditions from BLP are:

$$\begin{aligned} E\left[\xi_j(\theta)|X, W\right] &= 0 \\ E\left[w_j(\theta)|X, W\right] &= 0 \end{aligned}$$

- Instruments include Berry instruments and the product characteristics

15 Results

- Table 1 describes the product characteristics (average, sales-weighted)

- Table 2, purchaser characteristics.

- Minivans, higher income, bigger families than average, midage higher

- Table 3, minivans appear to be canabalizing station wagons

- Table 4, baseline demand estimates.

- Demand more elastic with IV, random coefficients and demographics

- Wald tests reject ols and IV for more flexible specifications

- Most coefficients make sense and are significant

- Marginal costs parameters Table 6.

- Most coefficients have anticipated sign

- Counter factual experiments- world with and without minivan.

- Equilibrium prices endogeneously adjust

- Table 7- prices of GM and Ford vehicles go down because of minivan through competitive effects.

- Prices of other Chrsyler products go up.
- Table 8- CV from removing minivan is very unrealistic in logit model.
- CV from random coefficients with CEX most reasonable.
- CV in logit model driven by large realizations of the preference shocks as in figure 1.
- Table 9, markup estimates.
- Also unrealistic under logit model since the price elasticity is biased towards zero.
- Table 10, markups on other vehicles seemed to have fallen while minivan margins remained stable.

- Table 11 and 12- Chrysler profits went up as a result of minivan.

- Ford and GM profits down.

- GM and Ford minivans bit into Chrysler's profits

- Firms ignore externalities on each other through canibalization of each others shares (e.g. not collusion).

- Chrysler's minivan seems to have differentiated itself enough to maintain profits and margins.

- Table 13- most of the benefits of the new goods go to consumers, not producers.