Discrete-Choice Models of Demand

In these lecture notes we present a framework for estimating demand in industries which are of particular interest for industrial organization. These are industries where

- There are a large number of competing products (large relative to number of distinct producers)
- Products are differentiated; no two products are exactly alike.
- Consumers make "discrete choices": that is, they typically choose only one of the competing products.
- Examples: soft drinks, pharmaceuticals, potato chips, cars, air travel, &etc.

Consumer Behavioral Model

- There are J alternatives in market, indexed by $j = 1, \ldots, J$. Each purchase occasion, each consumer i divides her income y_i on (at most) one of the alternatives, and on an "outside good":

$$\max_{j,z} U_i(x_j, z) \quad \text{s.t. } p_j + p_z z = y_i$$

  where

  - x_j are characteristics of brand j, and p_j the price
  - z is quantity of outside good, and p_z its price
  - outside good (denoted $j = 0$) is the non-purchase of any alternative (that is, spending entire income on other types of goods).

- Substitute in the budget constraint ($z = \frac{y-p_j}{p_z}$) to derive *conditional indirect utility functions* for each brand:

$$U^*_{ij}(x_j, p_j, p_z, y) = U_i(x_j, \frac{y_i - p_n}{p_z}).$$

If outside good is bought:

$$U^*_{i0}(p_z, y) = U_i(0, \frac{y_i - p_n}{p_z}).$$

- Consumer chooses the brand yielding the highest cond. indirect utility:

$$\max_j U^*_{ij}(x_j, p_j, p_z, y_i)$$

Econometric Model

- U^*_{ij} usually specified as sum of two parts:

$$U^*_{ij}(x_j, p_j, p_z, y_i) = V_{ij}(x_j, p_j, p_z, y_i) + \epsilon_{ij}$$

  ϵ_{ij} observed by agent i, not by researcher. It represents all else that affects consumers i's choosing product j, but is not observed by researcher.

- Because of the ϵ_{ij}'s, the product that consumer i chooses is random, from the researcher's point of view. Thus this model is known as the "random utility" model. (Daniel McFadden won the Nobel Prize in Economics for this, in 2000.)

- Specific assumptions about the ϵ_{ij}'s will determine consumer i's **choice probabilities**, which corresponds to her "demand function". Probability that consumer i buys brand j is:

$$D_{ij}(p_1 \dots p_J, p_z, y_i) = \text{Prob}\left\{\epsilon_{i0}, \dots, \epsilon_{iJ} : U^*_{ij} > U^*_{ij'} \text{ for } j' \neq j\right\}$$

Examples:

For now, consider the simple case where $J = 1$ (so that consumer i chooses either to buy good 1, or buy the "outside good, which is nothing at all).

With only two goods, consumer i chooses good 1 if

$$\begin{aligned} &V_{i0} + \epsilon_{i0} \leq V_{i1} + \epsilon_{i1} \\ \Leftrightarrow &V_{i0} - V_{i1} \leq \epsilon_{i1} - \epsilon_{i0}. \end{aligned} \tag{1}$$

Probit: If we assume

$$\eta_i \equiv \epsilon_{i1} - \epsilon_{i1} \sim N(0,1)$$

then we have the "probit" model. In this case, the choice probabilities are

$$P_{i1} = Pr(\eta_i \geq V_{i0} - V_{i1}) = 1 - \Phi(V_{i0} - V_{i1}).$$

In the above, $\Phi(x)$ is the cumulative distribution function (CDF) of a standard normal random variable, ie.

$$\Phi(x) = Pr(\eta_i < x).$$

Logit: If $(\epsilon_{ij},\ j = 0,1)$ are distributed *i.i.d.* type I extreme value across i, with CDF:

$$F(x) = \exp\left[-\exp\left(-\frac{x-\eta}{\mu}\right)\right] = Pr[\epsilon \leq x]$$

with $\eta = 0.577$ (Euler's constant), and the scale parameter (usually) $\mu = 1$, then we have the "logit" model. In this case, the choice probabilities are:

$$P_{i1} = \frac{\exp(V_{i1})}{\exp(V_{i0}) + \exp(V_{i1})}.$$

Multinomial Logit: One attractive feature of the logit model is that the choice probabilities scale up easily,[1] when we increase the number of products. For J products, the choice probabilities take the following "multinomial logit" form:

$$P_{ij} = \frac{\exp(V_{ij})}{\sum_{j'=0,\ldots,J} \exp(V_{ij'})}$$

Since this is a convenient form, the MNL model is often used in discrete-choice settings.

Problems with multinomial logit

[1] This is not true for the probit model.

- Restrictive implication of multinomial logit: "Odds ratio" (the ratio of two choice probabilities) between any two brands n, n' doesn't depend on number of alternatives available

$$\frac{P_{ij}}{P_{in'}} = \frac{\exp(V_{ij})}{\exp(V_{in'})}$$

regardless of which alternative brands $n'' \neq n' \neq n$ are available.

Consider:**Red bus/blue bus problem**

  - Assume that city has two transportation schemes: walk, and red bus. Consumer rides bus half the time, and walks to work half the time (ie. $P_{i,W} = P_{i,RB} = 0.5$). So odds ratio of walk/RB= 1.
  - Now consider introduction of third option: train. IIA implies that odds ratio between walk/red bus is still 1.

    This is behaviorally unrealistic: if train substitutes more with bus than walking, then new probs could be:

$$\begin{aligned} P_{i,W} &= 0.45 \\ P_{i,RB} &= 0.30 \\ P_{i,T} &= 0.25 \end{aligned}$$

    In this case, odds ratio walk/RB=1.5.
  - Now consider even more extreme case. What if third option were blue bus? IIA implies that odds ratio between walk/red bus would still be 1. This is very unrealistic: BB is perfect substitute for RB, so that new choice probabilities should be

$$\begin{aligned} P_{i,W} &= 0.50 \\ P_{i,RB} &= 0.25 \\ P_{i,BB} &= 0.25 \end{aligned}$$

    and odds ratio walk/RB=2!
  - So this is especially troubling if you want to use logit model to predict penetration of new products.

– What about using Logit model to predict elections?

Implication: invariant to introduction (or elimination) of some alternatives. **Independence of Irrelevant Alternatives**

- Since D_{ij} is demand function, IIA implies restrictive substitution patterns:

$$\varepsilon_{a,c} = \varepsilon_{b,c}, \text{ for all brands } a, b \neq c$$

where

$$\epsilon_{a,c} = \frac{\partial D_{ia}}{\partial p_c} \frac{p_c}{D_{ia}}$$

is the *cross-price elasticity of demand* between brands c and a. It is the percentage change in demand for product a, caused by a change in the price of product c.

If $V_{ij} = \beta_j + \alpha(y_i - p_j)$, then $\varepsilon_{a,c} = \alpha p_c D_c$, for all $c \neq a$: Price decrease in brand a attracts proportionate chunk of demand from all other brands. Unrealistic!

eg. Consider a price decrease in Cheerios. Naturally, demand for Cheerios would increase. IIA feature implies that demand for *all other brands* would decrease proportionately. This is unrealistic, because you expect that demand would decrease more for close substitutes of Cheerios (such as Honey Nut Cheerios), and less from non-substitutes (like Special K).

- Changes to logit framework to overcome IIA: *nested logit* model

Assume particular correlation structure among $(\epsilon_{i0}, \ldots, \epsilon_{ij})$. Within-nest brands are "closer substitutes" than across-nest brands.

(Diagram of demand structure from Goldberg paper)

Another problem: price endogeneity?

Note: in deriving all these examples, implicit assumption is that the distribution of the ϵ_{ij}'s are independent of the prices. This is analogous to assuming that prices are exogenous.

Case study: Trajtenberg (1989) study of demand for CAT scanners. Disturbing finding: coefficient on price is *positive*, implying that people prefer more expensive machines!

(Tables of results from Trajtenberg paper)

Possible explanation: quality differentials across products not adequately controlled for. In differentiated product markets, where each product is valued on the basis of its characteristics, brands with highly-desired characteristics (higher quality) may command higher prices. If any of these characteristics are not observed, and hence not controlled for, we can have endogeneity problems. ie. $E(p\epsilon) \neq 0$.

Estimation with aggregate market shares

Next we consider how to estimate demand functions in the presence of price endogeneity, and when the researcher only has access to *aggregate market shares*. This summarizes findings from Berry (1994).

Data structure: *cross-section* of market shares:

j	$\hat{s}_j$	p_j	X_1	X_2
A	25%	\$1.50	red	large
B	30%	\$2.00	blue	small
C	45%	\$2.50	green	large

Total market size: M
J brands

Try to estimate demand function from differences in market shares and prices across brands.

■■■

Derive market-level share expression from model of discrete-choice at the individual household level (i indexes household, j is brand):

$$U_{ij} = \underbrace{X_j\beta - \alpha p_j + \xi_j}_{\equiv \delta_j} + \epsilon_{ij}$$

where we call δ_j the "mean utility" for brand j (the part of brand j's utility which is common across all households i).

We need to estimate the model parameters (α, β).

■■■

Econometrician observes neither ξ_j or ϵ_{ij}, but household i observes both.

$\xi_1, \ldots, \xi_J$ are commonly interpreted as "unobserved product characteristics" or "unobserved quality". All else equal, consumers more willing to pay for brands for which ξ_j is high.

Important: ξ_j, as unobserved quality, is correlated with price p_j (and also potentially with characteristics X_j). It is the source of the endogeneity problem in this demand model.

Make multinomial logit assumption that $\epsilon_{ij} \sim \ iid$ TIEV, across consumers i and brands j.

Define choice indicators:

$$y_{ij} = \begin{cases} 1 & \text{if } i \text{ chooses brand } j \\ 0 & \text{otherwise} \end{cases}$$

Given these assumptions, choice probabilities take MN logit form:

$$Pr\left(y_{ij} = 1 | \beta, x_{j'}, \xi_{j'}, j' = 1, \ldots, J\right) = \frac{\exp\left(\delta_j\right)}{\sum_{j'=0}^{J} \exp\left(\delta_{j'}\right)}.$$

Aggregate market shares are (for $j = 0, \ldots, J$):

$$s_j = \frac{1}{M}\left[M \cdot Pr\left(y_{ij} = 1 | \beta, x_{j'}, \xi_{j'}, j' = 1, \ldots, J\right)\right] = \frac{\exp\left(\delta_j\right)}{\sum_{j'=0}^{J} \exp\left(\delta_{j'}\right)}$$
$$\equiv \tilde{s}_j\left(\delta_0, \ldots, \delta_J\right).$$

$\tilde{s}(\cdots)$ is the "predicted share" function, for fixed values of the parameters α and β, and the unobservables $\xi_1, \ldots, \xi_J$.

■■■

Estimation principle

- Data contains observed shares: denote by $\hat{s}_j, j = 1, \ldots, J$

  (Share of outside good is just $\hat{s}_0 = 1 - \sum_{j=1}^{J} \hat{s}_j$.)
- Model + parameters give you predicted shares: $\tilde{s}_j(\alpha, \beta, \xi_1, \ldots, \xi_J), j = 1, \ldots, J$
- Principle: Estimate parameters α, β by finding those values which "match" observed shares to predicted shares: find α, β so that $\tilde{s}_j(\alpha, \beta)$ is as close to $\hat{s}_j$ as possible, for $j = 1, \ldots, J$.

Berry (1994) suggests a clever IV-based estimation approach.

Assume there exist instruments Z so that

$$E\left(\xi Z\right) = 0 \tag{2}$$

Note that

$$\xi = \delta - X\beta_0 + \alpha_0 p$$

where (α_0, β_0) are the true, but unknown values of the model parameters. Hence, equation (2) can be written as

$$E[(\delta - X\beta_0 + \alpha_0 p)Z] = 0. \tag{3}$$

Sample version of this moment condition is

$$\frac{1}{J}\sum_{j=1}^{J}\left(\delta_j - X_j\beta + \alpha p_j\right) Z_j \equiv Q_j(\alpha,\beta).$$

We will estimate $(\alpha\beta)$ by minimizing

$$\min_{\alpha,\beta}[Q_J(\alpha,\beta)]^2. \tag{4}$$

Why does this work? As J gets large, by the law of large numbers, $Q_J(\alpha,\beta)$ converges to $E[(\delta - X\beta + \alpha p)Z]$. By equation (3), this is equal to zero at the true values (α_0,β_0). Hence, the (α,β) that minimize (4) should be close to (α_0,β_0). (And indeed, should converge to (α_0,β_0) as $J\rightarrow\infty$.)

Problem with estimating: we do not know δ_j! Berry suggest a *two-step approach*

First step: Inversion

- If we equate $\hat{s}_j$ to $\tilde{s}_j\left(\delta_0,\dots,\delta_J\right)$, for all j, we get a system of $J+1$ nonlinear equations in the $J+1$ unknowns $\delta_0,\dots,\delta_J$:

$$\begin{aligned}
\hat{s}_0 &= \tilde{s}_0\left(\delta_0,\dots,\delta_J\right)\\
\hat{s}_1 &= \tilde{s}_1\left(\delta_0,\dots,\delta_J\right)\\
\vdots\ & \quad\ \vdots\\
\hat{s}_J &= \tilde{s}_J\left(\delta_0,\dots,\delta_J\right)
\end{aligned}$$

- Note: the outside good is $j=0$. Since $1=\sum_{j=0}^{J}\hat{s}_j$ by construction, the equations are linear dependent. So you need to normalize $\delta_0=0$, and only use the J equations for $s_1,\dots,s_J$.

- Now you can "invert" this system of equations to solve for $\delta_1,\dots,\delta_J$ as a function of the observed $\hat{s}_0,\dots,\hat{s}_J$.

- Output from this step: $\hat{\delta}_j \equiv \delta_j\left(\hat{s}_0,\dots,\hat{s}_J\right),\ j=1,\dots,J$ (J numbers)

Second step: IV estimation

- Going back to definition of δ_j's:

$$\begin{aligned}\delta_1 &= X_1\beta - \alpha p_1 + \xi_1 \\ \vdots \quad & \quad \vdots \\ \delta_J &= X_J\beta - \alpha p_J + \xi_J\end{aligned}$$

- Now, using estimated $\hat{\delta}_j$'s, you can calculate $Q_J(\alpha, \beta)$

$$\frac{1}{J}\sum_{j=1}^{J}\left(\hat{\delta}_j - X_j\beta + \alpha p_j\right) Z_j$$

and solve for α, β as in equation (4) above.

■■■

The multonomial logit model yields a simple example of the inversion step.

MNL case: predicted share $\tilde{s}_j\left(\delta_1, \ldots, \delta_J\right) = \frac{\exp(\delta_j)}{1+\sum_{j'=1}^{J}\exp\left(\delta_{j'}\right)}$

The system of equations from matching actual to predicted shares is (note that here we have set $\delta_0 = 0$):

$$\begin{aligned}\hat{s}_0 &= \frac{1}{1+\sum_{j=1}^{J}\exp(\delta_j)} \\ \hat{s}_1 &= \frac{\exp(\delta_1)}{1+\sum_{j=1}^{J}\exp(\delta_j)} \\ \vdots \quad & \quad \vdots \\ \hat{s}_J &= \frac{\exp(\delta_J)}{1+\sum_{j=1}^{J}\exp(\delta_j)}.\end{aligned}$$

Taking logs, we get system of linear equations for δ_j's:

$$\begin{aligned}
\log \hat{s}_1 &= \delta_1 - \log(\text{denom}) \\
\vdots \quad & \quad \vdots \\
\log \hat{s}_J &= \delta_J - \log(\text{denom}) \\
\log \hat{s}_0 &= 0 - \log(\text{denom})
\end{aligned}$$

which yield

$$\delta_j = \log \hat{s}_j - \log \hat{s}_0, \quad j = 1, \ldots, J.$$

So in second step, run IV regression of

$$(\log \hat{s}_j - \log \hat{s}_0) = X_j\beta - \alpha p_j + \xi_j. \tag{5}$$

Eq. (5) is called a "logistic regression" by bio-statisticians, who use this logistic transformation to model "grouped" data. So in the simplest MNL logit, the estimation method can be described as "logistic IV regression".

See Berry paper for additional examples (nested logit, vertical differentiation).

■■■

What are appropriate instruments (Berry, p. 249)?

- Usual demand case: cost shifters. But since we have cross-sectional (across brands) data, we require instruments to very across brands in a market.
- Take the example of automobiles. In traditional approach, one natural cost shifter could be wages in Michigan.
- But here it doesn't work, because its the same across all car brands (specifically, if you ran 2SLS with wages in Michigan as the IV, first stage regression of price p_j on wage would yield the same predicted price for all brands).

- BLP exploit competition within market to derive instruments. They use IV's like: characteristics of cars of competing manufacturers. Intuition: oligopolistic competition makes firm j set p_j as a function of characteristics of cars produced by firms $i \neq j$ (e.g. GM's price for the Hum-Vee will depend on how closely substitutable a Jeep is with a Hum-Vee). However, characteristics of rival cars should not affect households' valuation of firm j's car.

- In multiproduct context, similar argument for using characteristics of all other cars produced by same manufacturer as IV.

- With panel dataset, where prices and market shares for same products are observed across many markets, could also use prices of product j in other markets as instrument for price of product j in market t (eg. Nevo (2001), Hausman (1996)).

■■■

0.1 Measuring market power: recovering markups

- Next, we show how demand estimates can be used to derive estimates of firms' markups (as in monopoly example from the beginning).
- From our demand estimation, we have estimated the demand function for brand j, which we denote as follows:

$$D^j\left(\underbrace{X_1,\ldots,X_J}_{\equiv\vec{X}};\underbrace{p_1,\ldots,p_J}_{\equiv\vec{p}};\underbrace{\xi_1,\ldots,\xi_J}_{\equiv\vec{\xi}}\right)$$

- Specify costs of producing brand j:

$$C^j\left(q_j\right)$$

  where q_j is total production of brand j.

- Then profits for brand j are:

$$\Pi_j = D^j\left(\vec{X},\vec{p},\vec{\xi}\right)p_j - C^j\left(D^j\left(\vec{X},\vec{p},\vec{\xi}\right)\right)$$

- For multiproduct firm: assume that firm k produces all brands $j \in \mathcal{K}$. Then its profits are

$$\tilde{\Pi}_k = \sum_{j\in\mathcal{K}}\Pi_j = \sum_{j\in\mathcal{K}}\left[D^j\left(\vec{X},\vec{p},\vec{\xi}\right)p_j - C^j\left(D^j\left(\vec{X},\vec{p},\vec{\xi}\right)\right)\right].$$

  Importantly, we assume that there are no (dis-)economies of scope, so that production costs are simply additive across car models, for a multiproduct firm.

  Aside Assume a firm produces two items (call it "1" and "2"). Let $C(q_1,q_2)$ denotes the total cost function of producing q_1 units of good 1 and q_2 units of good 2. Let $C_1(q_1)$ be the total cost function of just producing good 1; and $C_2(q_2)$ be total cost function of just producing good 2.

  - Economies of scope: $C(q_1,q_2) < C_1(q_1) + C_2(q_2)$
  - Diseconomies of scope: $C(q_1,q_2) > C_1(q_1) + C_2(q_2)$

  - No economies of diseconomies of scope: $C(q_1, q_2) = C_1(q_1) + C_2(q_2)$

- In order to proceed, we need to assume a particular model of oligopolistic competition.

  One common assumption is *Bertrand (price) competition.* (Note that because firms produce differentiated products, Bertrand solution does not result in marginal cost pricing.)

  That is, firm k chooses prices p_j, $j \in \mathcal{K}$, to maximize $\tilde{\Pi}_k$ (defined above)

- Under price competition, equilibrium prices are characterized by J equations (which are the J pricing first-order conditions for the J brands):

$$\frac{\partial \tilde{\Pi}_k}{\partial p_j} = 0, \ \forall j \in \mathcal{K}, \ \forall k$$
$$\Leftrightarrow D^j + \sum_{j' \in \mathcal{K}} \frac{\partial D^{j'}}{\partial p_j} \left(p_{j'} - C_1^{j'}|_{q_{j'}=D^{j'}} \right) = 0$$

  where C_1^j denotes the derivative of C^j with respect to argument (which is the marginal cost function).

- Note that because we have already estimated the demand side, the demand functions D^j, $j = 1, \ldots, J$ and full set of demand slopes $\frac{\partial D^{j'}}{\partial p_j}$, $\forall j, j' = 1, \ldots, J$ can be calculated.

  Indeed, for the MNL model, and assuming that $V_j = X_j\beta - \alpha p_j$, then

$$\frac{\partial D^{j'}}{\partial p_j} = \begin{cases} -\alpha D^j(1 - D^j) & \text{if } j = j' \\ \alpha D^{j'} D^j & \text{if } j \neq j' \end{cases}$$

  Hence, from these J equations, we can solve for the J margins $p_j - C_1^j$. In fact, the system of equations is linear, so the solution of the marginal costs C_1^j is just

$$\vec{c} = \vec{p} + (\Delta D)^{-1} \vec{D}$$

  where c and D denote the J-vector of marginal costs and demands, and the derivative matrix ΔD is a $J \times J$ matrix where

$$\Delta D_{(i,j)} = \begin{cases} \frac{\partial D^i}{\partial p_j} & \text{if models } (i,j) \text{ produced by the same firm} \\ 0 & \textit{otherwise.} \end{cases}$$

The markups measures can then be obtained as $\frac{p_j - C_1^j}{p_j}$.

This is the oligopolistic equivalent of using the "inverse-elasticity" condition to calculate a monopolist's market power.

Random coefficient logit model In recent applications of discrete-choice demand models, focus on a more complicated version of MNL.

For agent i, utility from brand j is:

$$U_{ij}^* = X_j' \beta_i - \alpha_i p_j + \xi_j + \epsilon_{ij}$$

(coefficients are agent-specific).

ϵ_{ij} remain TIEV, so underlying model is logit. However now the coefficients β_i and α_i differ across households i.

Assume that β_i and α_i are distributed across consumers according to a normal distribution $N([\bar{\alpha}, \bar{\beta}]', \Sigma)$.

Then you can write the above as:

$$U_{ij}^* = \underbrace{X_j' \bar{\beta} - \bar{\alpha} p_j + \xi_j}_{\delta_j} + X_j'(\beta_i - \bar{\beta}) - (\alpha_i - \bar{\alpha}) p_j + \epsilon_{ij}$$

where δ_j denotes mean utility, as before.

Hence, in this model, aggregate market share is not equal to consumer-level choice probability. In fact, aggregate market share is *integral* of consumer-level choice probabilities.

Consumer-level choice probability:

$$P_{ij} = \frac{\exp(\delta_j + X_j'(\beta_i - \bar{\beta}) - (\alpha_i - \bar{\alpha}) p_j)}{1 + \sum_{j'=1}^{J} \exp(\delta_{j'} X_{j'}'(\beta_i - \bar{\beta}) - (\alpha_i - \bar{\alpha}) p_{j'})}$$

Then aggregate market share is

$$\begin{aligned} D_j &= \int P_{ij} dF(\alpha_i, \beta_i) \\ &= \frac{\exp(\delta_j + X_j'(\beta_i - \bar{\beta}) - (\alpha_i - \bar{\alpha})p_j)}{1 + \sum_{j'=1}^{J} \exp(\delta_{j'} X_{j'}'(\beta_i - \bar{\beta}) - (\alpha_i - \bar{\alpha})p_{j'})} dF(\alpha_i, \beta_i). \end{aligned}$$

It turns out estimation of this model (and accommodating endogeneity) is quite complicated and involved, and we will not discuss it here.

1 Applications

Applications of this methodology have been voluminous. Here discuss just a few.

1. evaluation of VERs In Berry, Levinsohn, and Pakes (1999), this methodology is applied to evaluate the effects of voluntary export restraints (VERs). These were voluntary quotas that the Japanese auto manufacturers abided by which restricted their exports to the United States during the 1980's.

The VERs do not affect the demand-side, but only the supply-side. Namely, firm profits are given by:

$$\pi_k = \sum_{j \in \mathcal{K}} (p_j - c_j - \lambda VER_k) D^j.$$

In the above, VER_k are dummy variables for whether firm k is subject to VER (so whether firm k is Japanese firm). VER is modelled as an "implicit tax", with $\lambda \geq 0$ functioning as a per-unit tax: if $\lambda = 0$, then the VER has no effect on behavior, while $\lambda > 0$ implies that VER is having an effect similar to increase in marginal cost c_j. The coefficient λ is an additional parameter to be estimated, on the supply-side.

Results (effects of VER on firm profits and consumer welfare)

2. Welfare from new goods, and merger evaluation After cost function parameters γ are estimated, you can simulate equilibrium prices under alternative market structures, such as mergers, or entry (or exit) of firms or goods. These counterfactual prices are valid assuming that consumer preferences and firms' cost functions don't change as market structures change. Petrin (2002) presents consumer welfare benefits from introduction of the minivan, and Nevo (2001) presents merger simulation results for the ready-to-eat cereal industry.

3. Geographic differentiation In our description of BLP model, we assume that all consumer heterogeneity is unobserved. Some models have considered types of consumer heterogeneity where the marginal distribution of the heterogeneity in the population is observed. In BLP's original paper, they include household income in the utility functions, and integrate out over the population income distribution (from the Current Population Survey) in simulating the predicted market shares.

Another important example of this type of obbserved consumer heterogeneity is consumers' location. The idea is that the products are geographically differentiated, so that consumers might prefer choices which are located closer to their home. Assume you want to model competition among movie theaters, as in Davis (2006). The utility of consumer i from theater j is:

$$U_{ij} = -\alpha p_j + \beta(L_i - L_j) + \xi_j + \epsilon_{ij}$$

where $(L_i - L_j)$ denotes the geographic distance between the locations of consumer I and theater j. The predicted market shares for each theater can be calculated by integrating out over the marginal empirical population density (ie. integrating over the distribution of L_i). See also Thomadsen (2005) for a model of the fast-food industry, and Houde (2006) for retail gasoline markets. The latter paper is noteworthy because instead of integrating over the marginal distribution of where people live, Houde integrates over the distribution of commuting routes. He argues that consumers are probably more sensitive to a gasoline station's location relative to their driving routes, rather than relative to their homes.

2 Panel data and fixed-effects regression

Define: *panel dataset* is a dataset which contains multiple observations for the same unit, over time. Let i index units, and t index time periods.

Suppose you are interested in estimating the relationship:

$$Y_{it} = \alpha + \beta X_{it} + \epsilon_{it}.$$

Having panel data allow you to control for *unit-specific unobservables* γ_i as well as *time-specific* unobservables δ_t:

$$Y_{it} = \alpha + \beta X_{it} + \gamma_i + \delta_t + \epsilon_{it}.$$

References

Berry, S. (1994): "Estimating Discrete Choice Models of Product Differentiation," *RAND Journal of Economics*, 25, 242–262.

Berry, S., J. Levinsohn, and A. Pakes (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, 63, 841–890.

——— (1999): "Voluntary Export Restraints on Automobiles: Evaluating a Strategic Trade Policy," *American Economic Review*, 89, 400–430.

Davis, P. (2006): "Spatial Competition in Retail Markets: Movie Theaters," *RAND Journal of Economics*, pp. 964–982.

Hausman, J. (1996): "Valuation of New Goods under Perfect and Imperfect Competition," in *The Economics of New Goods*, ed. by T. Bresnahan, and R. Gordon, pp. 209–237. University of Chicago Press.

Houde, J. (2006): "Spatial Differentiation in Retail Markets for Gasoline," Manuscript, University of Wisconsin.

Nevo, A. (2001): "Measuring Market Power in the Ready-to-eat Cereals Industry," *Econometrica*, 69, 307–342.

Petrin, A. (2002): "Quantifying the Benefits of New Products: the Case of the Minivan," *Journal of Political Economy*, 110, 705–729.

Thomadsen, R. (2005): "The Effect of Ownership Structure on Prices in Geographically Differentiated Industries," *RAND Journal of Economics*, pp. 908–929.

Trajtenberg, M. (1989): "The Welfare Analysis of Product Innovations, with an Application to Computed Tomography Scanners," *Journal of Political Economy*, 97(2), 444–479.