Differentiated products

Meredith Fowlie
UC Berkeley

October 2021

life
Honey Graham Oh's
Always a GREAT VALUE
CAP'N CRUNCH
CRUNCH-ATIZE ME CAP'N!
Incredible Berry Crunch!
Lucky Charms
CAP'N CRUNCH'S PEANUT BUTTER CRUNCH
Berry Lucky Charms
NEW LOOK, SAME GREAT TASTE!
OATMEAL CRISP
Crunchy Almond
Honey Nut Chex
NEW cinnamon STARS
Cinnamon Toast Crunch
Whole Grain
CRISP
Reese's Puffs
Kid-Tested
KIX
Mother-Approved!
Total
Whole Grain
Kellogg's
HONEY SMACKS
FROSTED Mini-Wheats Bite Size
Trix
FREE GAME RENTALS
FROSTED Mini-Wheats Strawberry Delight
CORN FLAKES
Free MOVIE
NEW!
Eggo cereal
CINNAMON TOAST
FAMILY SIZE

Motivation

Economists often interested in estimating demand for differentiated products in markets with many choices. EEE examples/applications:

- Vehicles (e.g. impacts of fuel efficiency standards? EV mandates?)
- Transportation modes (e.g. welfare impacts of improved public transit options, bike lanes?)
- Recreation area choices (e.g. valuation of improved amenities)
- Appliances (e.g. valuation of energy efficiency attributes)
- Neighborhood choice (e.g. valuing proximity to public transport, parks, etc.)

Curse of dimensionality in markets with many varieties

We are often interested in how consumers substitute between choices/products. Consider estimating an unrestricted demand system for J differentiated products:

$$
\begin{aligned}
Q_j &= X_j'\beta + \sum_k \alpha_{jk} p_{jk} + u_{jk}, \\
|\alpha| &= JxJ
\end{aligned}
$$

- Estimating an unrestricted demand system involves *at least JxJ* parameters.
- Even if we impose reasonable restrictions (such as a symmetric $|\alpha|$ matrix) number of parameters to estimate unwieldy.
- Also difficult to accommodate heterogeneity in tastes (a critical consideration in markets for differentiated products!)

Solution?

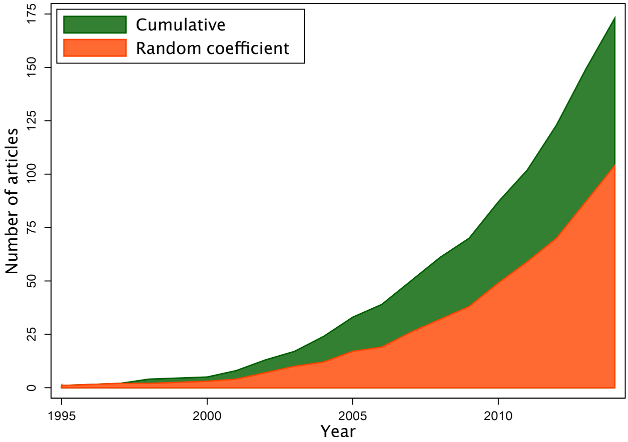
- Project differentiated products onto characteristic/attribute space.
- Projection into finite attribute space can an allow us to derive more tractable demand systems in markets with many differentiated products/choices.
- The underlying idea is related to our earlier discussion of hedonic models: Consumers' preferences for different product characteristics define consumers' preferences/demand for product varieties.
- But today's emphasis is more structural. Commit to a specific utility function in order to estimate underlying taste parameters (versus MWTP).

Cue Berry, Levinsohn, and Pakes!

- In a series of papers written by Steven Berry along with co-authors Jim Levinsohn, and Ariel Pakes, the CL and RCL frameworks are augmented and extended so as to address (or at least acknowledge) these challenges.
- They develop an empirically tractable framework which enables one (in principle) to obtain estimates of demand and cost parameters for a class of oligopolistic differentiated product markets using only product-level and aggregate consumer-level data.
- They use data on unit sales, list prices, vehicle attributes, and market demographics to draw inferences about consumer preferences and automobile manufacturers' margins.
- These are a very rich set of papers which inspired a huge applied literature that spans numerous fields (industrial organization, health, development, labor, environment).

20 years of BLP(1995)

Over 3,300 citations and 175 applications



Over the past two decades, there has been an explosion of empirical work using models of differentiated product markets that build on BLP(1995)

BLP

In this lecture, we will be focusing on three BLP innovations in particular:

1. Estimating models of consumer-level demand using aggregate (i.e. market level) versus disaggregated (i.e. household or customer level) data (hereby generalizing the use of simulation methods)
2. Demonstrating a tractable means of dealing with endogenous covariates (this is an extension of the Berry transformation we introduced last time).
3. Nesting the random utility model within a larger structural model of imperfect competition in order to obtain consistent estimation of underlying structural parameters.

Working example: Demand for new vehicles

Some assumptions we'll invoke(subsequent work has sought to release some of these assumptions):

- Firms sell new cars directly to consumers.
- Firms do not price discriminate.
- Consumers know the prices and attributes of all new cars on the market.
- There are no dynamic considerations for either firms or consumers. In particular, consumers do not trade off prices and product attributes today with those in the future.
- All non-price product attributes are assumed to be exogenous- vehicle price endogenous.
- Consumers purchase at most one car per household

BLP: Basic Set-up

- Seminal 1995 paper uses market-level (versus consumer-level) data on car sales 1971-1990.
- Authors do not observe individual vehicle choices. Instead they observe market shares, listed prices, demographics, some attributes, and product/market level cost shifters, but not individual choices or firm costs.
- How the heck can they estimate a discrete choice model with aggregage (market-level) data?

BLP: Basic Set-up

- Seminal 1995 paper uses market-level (versus consumer-level) data on car sales 1971-1990.
- Authors do not observe individual vehicle choices. Instead they observe market shares, listed prices, demographics, some attributes, and product/market level cost shifters, but not individual choices or firm costs.
- How the heck can they estimate a discrete choice model with aggregage (market-level) data?
- Recall from last lecture: There is an intuitive conceptual mapping between choice probabilities and market shares. We'll elaborate on that idea in a more complicated setting.

Notation

Authors have twenty years of data covering 2217 (that's a lot) new car models:

- Products/vehicle models are indexed by $j = 0...J_t$.
- Markets are indexed by $t = 1...T$.
- Individuals are indexed by $n = 1....N_t$.

How to define a market?

- When we are using aggregate data (versus consumer-level data) to estimate these demand models, estimation based on matching aggregate market shares to estimated choice probabilities.
- Market shares are constructed as sales divided by a hypothesized number of potential sales.
- In general, we should think about product attributes (including prices) varying across markets, but not across consumers within a given market.
- We'll return to this design detail when we think about empirical estimation of market shares....

Choice set?

- Assume that j indexes mutually exclusive, exhaustive, and finite set of choices.
- In general, consumers are assumed to choose the option that yields the highest utility.
- Here we assume the consumer chooses at most one of the differentiated products.
- Outside good? BLP assume consumers choose between $J + 1$ options, where $j = 0$ is usually reserved for the outside option.
- The inclusion of the outside good allows us to use these models to study aggregate demand because we do not condition on purchasing a new car.

RUM (BLP-style)

- Mixed logit provides the foundation for the demand model.
- Consumer n derives utility from choice j in market t.
- Assume that the consumer chooses the option that maximizes utility.
- Allow taste parameters to vary with observable demographics.
- Utility is a function of the expenditures on other goods and services and the attributes of the differentiated good:

$$u_{njt} = \alpha \ln(y_n - p_{jt}) + \sum_k x_{jtk} \beta_{nk} + \xi_{jt} + \varepsilon_{njt}. \tag{1}$$

What's going on with income?

$$u_{njt} = \alpha_n \ln(y_n - p_{jt}) + \sum_k x_{jtk}\beta_{nk} + \xi_{jt} + \varepsilon_{njt}. \tag{2}$$

- Income y_n enters so that we can interpret latent utility as a conditional indirect utility function (i.e. consumer's maximized utility when faced with prices p and income y).
- If we assume quasi-linear utility (free of wealth effects), income enters as $\alpha_n(y_n - p_{jt})$.
- BLP assume a Cobb-Douglas utililty function which yields an indirect utility function that includes $\ln(y_n - p_{jt})$.
- The α is an unknown (to be estimated) marginal utility of income.

Random utility model (BLP-style)

$$u_{njt} = \alpha \ln(y_n - p_{jt}) + \sum_k x_{jtk}\beta_{nk} + \xi_{jt} + \varepsilon_{njt}. \tag{3}$$

- The price of product j in market t is p_{jt}. Assume all consumers in market t are offered the same price.
- The income of consumer n is y_n.
- Let k denote the observable non-price attributes.
- The X_{jt} matrix includes observable non-price product attributes that can vary across markets and products.
- The standard specification assumes all consumers in a market face the same product characteristics and the same prices. Subsequent work has released this assumption (see Langer, 2011).

What's that ξ_{jt} doing there?

$$u_{njt} = \alpha \ln(y_n - p_{jt}) + \sum_k x_{jtk}\beta_{nk} + \xi_{jt} + \varepsilon_{njt}. \tag{4}$$

- The ξ_{jt} is a structural error term.
- Captures average valuation of attributes and quality that are observed by consumers but are not observable to the researcher (and thus not explicitly represented in the model).
- Mechanically, these parameters work to ensure that the market shares predicted by the econometric model match the 'observed' market shares.
- BLP assume that the ξ_{jt} are 'mean independent' of the observed attributes of new cars. We'll come back to this...

It wouldn't be logit without ε_{njt}

$$u_{njt} = \alpha \ln(y_n - p_{jt}) + \sum_k x_{jtk}\beta_{nk} + \xi_{jt} + \varepsilon_{njt}. \tag{5}$$

- ε_{njt} is a white noise (by assumption) disturbance term.
- For convenient error partitioning, this residual is assumed to be distributed *iid* EV.
- Having included the ξ_{jt} to control the average effect of unobserved vehicle attributes, and allowing the preference coefficients to vary across consumers, there is presumably little variation left to be captured by this residual.
- This error term could capture measurement error, optimization error, etc.

The outside option

The specification of the demand system is completed by the outside option. The most standard approach involves setting the $j = 0$ parameters to zero:

$$u_{n0t} = \alpha \ln(y_n) + \beta_{n0} + \xi_{0t} + \varepsilon_{n0t}.$$

- Utility derived from the outside good is not identified.
- Standard practice is to set $\beta_{i0} = \xi_{0t} = 0$. The $\alpha \ln(y_n)$ drops out because it is common to all choices.
- In subsequent work (e.g. Berry et al, 2003), the specification of this outside option is defined to be a parametric function of observed household attributes.

How to model preference heterogeneity?

- In the context of demand models for differentiated products, it's important to accommodate preference heterogeneity (if preferences were homogeneous, why are product offerings differentiated?).
- In most cases, we will expect preferences to be systematically related to factors we can observe (e.g. gender, income, dental health) as well as factors we cannot observe.
- Allowing preference parameters to vary systematically with observables reduces our reliance on the (often arbitrary) parametric assumptions we impose on the unobservable random components.

How to model preference heterogeneity?

Decompose taste parameters into a deterministic and a stochastic component:

$$\begin{aligned} \alpha_n &= \alpha + \eta_{n\alpha} \\ \beta_{nk} &= \beta_k + \sum_r \mu_{kr} d_{nr} + \eta_{nk} \\ & \eta_n | d_n \tilde{\ } N(0, \Sigma) \\ & d_n \tilde{\ } f(d_n) \end{aligned}$$

- d_n is a vector of consumer demographics. For expositional simplicity, I assume that all preference parameters are allowed to vary systematically with all household characteristics.
- Let r index the consumer characteristics we observe: $d_i = \{d_{i1} ... d_{ir}\}$.
- Demographic variables typically normalized to have mean zero so the β can be interpreted as average values.

How to model preference heterogeneity?

Decompose taste parameters into a deterministic and a stochastic component:

$$
\begin{aligned}
\alpha_n &= \alpha + \eta_{n\alpha} \\
\beta_{nk} &= \beta_k + \sum_r \mu_{kr} d_{nr} + \eta_{nk} \\
&\eta_n | d_n \tilde{} N(0, \Sigma) \\
&d_n \tilde{} f(d_n)
\end{aligned}
$$

- If we are estimating the model using market-level data, we do not observe d_{nr} at the individual level. So we use an estimate of the distribution of these variables in the market-level population.
- Let $f(d_{ir})$ be the distribution of the d_r in the population.

How to model preference heterogeneity?

Decompose taste parameters into a deterministic and a stochastic component:

$$
\begin{aligned}
\alpha_n &= \alpha + \eta_{n\alpha} \\
\beta_{nk} &= \beta_k + \sum_r \mu_{kr} d_{nr} + \eta_{nk} \\
& \eta_n | d_n \tilde{} N(0, \Sigma) \\
& d_n \tilde{} f(d_n)
\end{aligned}
$$

- The η_{ik} capture the effects of unobserved preference variation that we assume is randomly distributed in the population.
- Here I assume a normal distribution- other distributional assumptions are readily substituted.
- Because taste parameters vary systematically with demographics, we capture more of the variation in tastes in this systematic component, which reduces our reliance on the parametric assumptions governing η_{ik} .

How to model preference heterogeneity?

Decompose taste parameters into a deterministic and a stochastic component:

$$
\begin{aligned}
\alpha_n &= \alpha + \eta_{n\alpha} \\
\beta_{nk} &= \beta_k + \sum_r \mu_{kr} d_{nr} + \eta_{nk} \\
&\quad \eta_n | d_n \tilde{\ } N(0, \Sigma) \\
&\quad d_n \tilde{\ } f(d_n)
\end{aligned}
$$

- As noted above, BLP assume an indirect utility function derived from Cobb-Douglas utility.
- A first-order Taylor series approximation to this Cobb-Douglas implies that α_n is inversely proportional to income. So α_i sometimes modeled as $\frac{\alpha}{y_n}$.

Decomposition of indirect utility

Taken together, indirect utility can be expressed as a sum of four terms:

$$
\begin{aligned}
u_{njt} &= \underbrace{\alpha p_{jt} + X_{jt}\beta + \xi_{jt}}_{} + \underbrace{(\sum_r \mu_{kr} d_{nr} + \eta_{nk})' X_{jt} - (\mu_\alpha + \eta_{i\alpha})\frac{p_{jt}}{y_n}}_{} + \varepsilon_{njt} \\
&= \delta_{jt}(X_{jt}, p_{jt}, \xi_{jt}; \theta_1) + \upsilon_{njt}(X_{jt}.p_{jt}, \eta_n, d_{nr}; \theta_2) + \varepsilon_{njt} \\
&\quad \eta_n | d_n \tilde{} N(0, \Sigma) \\
&\quad d_n \tilde{} f(d_n) \\
&\quad \varepsilon_{njt} \tilde{} iid\ EV1
\end{aligned}
$$

- $\alpha p_{jt} + X_{jt}\beta + \xi_{jt}$ captures the mean utility term that is common to all consumers in market t.
- The model-specific mean utlility components are summarized as δ_{jt}.

Decomposition of indirect utility

Taken together, indirect utility can be expressed as a sum of four terms:

$$
\begin{aligned}
u_{njt} &= \underbrace{\alpha p_{jt} + X_{jt}\beta + \xi_{jt}} + \underbrace{(\sum_r \mu_{kr} d_{nr} + \eta_{nk})' X_{jt} - (\mu_\alpha + \eta_{i\alpha}) \frac{p_{jt}}{y_n}} + \varepsilon_{njt} \\
&= \delta_{jt}(X_{jt}, p_{jt}, \xi_{jt}; \theta_1) + \upsilon_{njt}(X_{jt}.p_{jt}, \eta_n, d_{nr}; \theta_2) + \varepsilon_{njt} \\
& \quad \eta_n | d_n \tilde{} N(0, \Sigma) \\
& \quad d_n \tilde{} f(d_n) \\
& \quad \varepsilon_{njt} \tilde{} iid\ EV1
\end{aligned}
$$

- The second element captures the individual-specific deviation from δ_{jt}.
- The ε_{ijt}, is the stochastic error term that is $\tilde{}$*iid* EV by assumption.

$$u_{njt} = \delta_j + (\beta_n - \overline{\beta}) X_{jt} - (\alpha_n - \overline{\alpha}) \frac{p_{jt}}{y_n} + \varepsilon_{njt}$$

Holy Notation Batman!

The convention is to divide the parameters to be estimated into two categories:

Linear parameters

- The parameter vector $\{\alpha, \beta\} = \theta_1$.
- These parameters are subsumed in the product specific constants δ.

Non-linear parameters

- $\{\mu, \Sigma\} = \theta_2$.
- If the η_{ik} have independent standard normal distributions, the μ_r capture the average deviation from the mean preference parameters as a function of the observable customer attribute d_r.
- The parameters in Σ capture the standard deviations of the contributions of unmeasured customer attributes to the deviation in the customer specific β parameters from the mean.

Unconditional probabilities

Once you have made your distributional assumptions about how the η_i vary in the population, you can write down the expression for the unconditional choice probabilities:

$$P_{nit} = \int \frac{\exp(\delta_{it} + \upsilon_{nit})}{1 + \sum \exp(\delta_{jt} + \upsilon_{njt})} f(\theta_2)$$

- To simplify notation, use υ_{njt} to stand-in for $(\beta_n - \overline{\beta})X_{jt} - (\alpha_n - \overline{\alpha})\frac{p_{jt}}{y_n}$
- The 1 in the denominator comes from setting the utility from the outside option to be zero.
- These choice probabilities look familiar!! This is just a souped-up mixed logit

Wrinkles and complications

In differentiated products markets with many products, it will often be the case that we will run into one or more of the following issues:

1. Often, we do not have individual choice data. More common to observe aggregate (i.e. market-level) data and the distribution of demographic characteristics in the market-specific populations (from the census, for example).
2. In most differentiated product markets, unobserved attributes introduce endogeneity concerns.
3. Estimating an RCL model with a large number of constants can be difficult using standard, gradient-based methods.

BLP

In this lecture, we will be focusing on three BLP innovations in particular:

1. Estimating models of consumer-level demand using aggregate (i.e. market level) versus disaggregated (i.e. household or customer level) data (hereby generalizing the use of simulation methods)
2. Demonstrating a tractable means of dealing with endogenous covariates (this is an extension of the Berry transformation we introduced last time).
3. Nesting the random utility model within a larger structural model of imperfect competition in order to obtain consistent estimation of underlying structural parameters.

Estimating a model of consumer-level demand using market -level data?

- The basic intuition: Market-level demand/market shares are simply the sum of individual consumer choices. So there is an intuitive mapping between choice probabilities and market shares.
- Identification relies on observable variation across markets (in the choice set faced by consumers, consumer attributes, and market shares).
- The authors also must assume that the distribution of consumers' underlying tastes, conditional on an observed distribution of consumer characteristics, is invariant across markets.

Market-level data

Estimation intuition: Choose parameter values that minimize the distance between observed and predicted *market shares*.

- Data: Combine observed data on market-specific product attributes (including price data), observed market shares, and data on how individual-level characteristics, d_i are distributed in each market.
- Assume each consumer purchases one unit of the good that yields the highest utility.
- If the outside good is not included, the model can be used to predict changes in market shares, but cannot be used to predict changes in total quantity demanded (because the number of consumers is effectively held fixed). to estimate overall market size. Not always straightforward!

Recall the Berry transformation (simple CL version)

Formalize the relationship betweein consumer-level choice probabilities and market shares. To illustrate the basic intuition, recall our CL example where $\mu = \Sigma = 0$.

$$\begin{aligned} U_{njt} &= \delta_{jt} + \varepsilon_{njt} \\ \Pr(U_{nit} &\geq U_{njt} \forall j) = \frac{\exp(\delta_{it})}{1 + \sum_j \exp(\delta_{jt})}, \end{aligned}$$

In this simple case, the estimated market share of product j is simply the predicted choice probability:

$$\widehat{s}_{it}(\delta_{it}, \mu = \Sigma = 0) = \Pr(U_{ni} \geq U_{nj} \forall j) = \frac{\exp(\delta_i)}{1 + \sum_j \exp(\delta_j)}. \tag{6}$$

How do we estimate this using only market level data?

Market-level data

Define market share equations $\widehat{s}_{it}(\delta_{it})$:

$$\widehat{s}_0 = \frac{1}{1 + \sum_j \exp(\delta_j)} \tag{7}$$

$$\widehat{s}_1 = \frac{\exp(\delta_1)}{1 + \sum_j \exp(\delta_j)} \tag{8}$$

$$\vdots \tag{9}$$

$$\widehat{s}_J = \frac{\exp(\delta_J)}{1 + \sum_j \exp(\delta_j)} \tag{10}$$

With these share equations in hand, we can in principle solve for the δ_{jt} that fit the observed shares exactly.

Invert the share equations

In the CL model, this can be done analytically. Taking logs:

$$\log \widehat{s}_0 = -\log(1 + \sum_j \exp(\delta_j)) \tag{11}$$

$$\log \widehat{s}_1 = \delta_1 - \log(1 + \sum_j \exp(\delta_j)) \tag{12}$$

$$\vdots \tag{13}$$

$$\log \widehat{s}_J = \delta_J - \log(1 + \sum_j \exp(\delta_j)) \tag{14}$$

This yields:

$$\log \widehat{s}_j - \log \widehat{s}_0 = \delta_j = X_j'\beta + \alpha p_j + \xi_j \tag{15}$$

The key intuition: There exists a unique δ such that the predicted shares equal actual (observed) shares.

Allow for heterogeneous preferences...

- Preference heterogeneity enters into the CL model only through the iid error term - which implies unrealistic substitution patterns.
- We want to allow for heterogeneous preferences in the population. This will, among other things, capture more realistic substitution patterns.
- To accommodate this, each (simulated) individual manifests as a bundle of observable characteristics d_i and random utility components η_i and product specific shocks ε_i.
- The overall market share of product j in market t is found by integrating the choices of each consumer across the distribution of consumers in the market.

Share equations when preferences vary in the population

We can estimate the aggregate share of the market t claimed by product j in terms of (shorthand notation a little sloppy for ease of exposition):

$$\begin{aligned}
\widehat{s}_{it}(\delta_{it}) &= \int\int \frac{\exp(X_{it}\beta_n - \alpha_n p_{it} + \xi_i)}{1 + \sum_j \exp(X_{jt}\beta_n - \alpha_n p_{jt} + \xi_j)} df(\alpha_n, \beta_n). \\
&= \int\int \frac{\exp(\delta_{jt} + X_{jt}(\beta_n - \overline{\beta}) - (\alpha_n - \overline{\alpha_n})p_{it})}{1 + \sum_j \exp(\delta_{jt} + X_{jt}(\beta_n - \overline{\beta}) - (\alpha_n - \overline{\alpha})p_{jt})} df(\alpha_n, \beta_n).
\end{aligned}$$

Now the predicted market shares are a function not only of the mean utilities δ_{jt}, but also the unknown Non-linear parameter estimates Θ_2

Share equations meet random coefficients

$$\widehat{s}_{it}(\delta_{it}) = \int\int \frac{\exp(\delta_{jt} + X_{jt}(\beta_n - \overline{\beta}) - (\alpha_n - \overline{\alpha_n})p_{it})}{1 + \sum_j \exp(\delta_{jt} + X_{jt}(\beta_n - \overline{\beta}) - (\alpha_n - \overline{\alpha})p_{jt})} df(\alpha_n, \beta_{in}).$$

- To estimate this we need: estimated market shares (LHS), prices, non-price attributes, consumer demographics.
- The δ_{jt} absorb product characteristics that are common across consumers (and markets if δ_j).
- Variation in product attributes across markets helps identify substitution patterns.

Estimation overview

- Collect data on product attributes X and market shares y.
- Specify latent utility function (which includes parametric assumptions about how random taste parameters are distributed in the population.
- Simulate drawing from these assumed distributions (generate a set of R draws) using an initial guess for the non-liner parameters.
- Estimate the implied market shares given these draws and an initial guess of the mean utilities δ and an initial guess of the non-linear parameters θ_2.
- Calibrate the simulated log likelihood function (SLL) or moment conditions and objective function (GMM).
- Use numerical optimization algorithms to identify the estimates of non-linear parameters and the δ parameters.
- Once converged, step outside and use estimated δ to recover the linear parameters.

Estimation via simulation

$$\widehat{s}_{it}(\delta_{it};\Theta_2) = \int\int \frac{\exp(\delta_{jt} + X_{jt}(\beta_n - \overline{\beta}) - (\alpha_n - \overline{\alpha_n})p_{it})}{1 + \sum_j \exp(\delta_{jt} + X_{jt}(\beta_n - \overline{\beta}) - (\alpha_n - \overline{\alpha})p_{jt})} df(\alpha_n, \beta_{in}).$$

- Given distributional assumptions, we can predict the market share of each product in each market as a function of the observed product characteristics, prices, random tastes, etc.
- To simulate this multi-dimensional integral, we are taking a weighted sum across choice probabilities of all individuals, where the weights are given by:
    - The probability distribution of the α_n and β_n in the population.
    - The distributions of the α_n and β_n in the population are, in turn, determined by the distribution of the η_{nk} and d_{nr} in the population.
- Non-linear parameter estimates Θ_2 are those that minimize the distance between estimated and observed market shares.

This gets a little easier with consumer-level data: 'Micro-BLP'

- BLP-type models can also be estimated using individual-level data.
- Now we search for the parameter values that best rationalize observed consumer choices versus estimated market shares.
- Matching consumer-level attribute data to customer choices can help us to more precisely characterize the relationship between observable customer characteristics and substitution patterns.

BLP

In this lecture, we will be focusing on three BLP innovations in particular:

1. Estimating models of consumer-level demand using aggregate (i.e. market level) versus disaggregated (i.e. household or customer level) data (hereby generalizing the use of simulation methods)
2. Demonstrating a tractable means of dealing with endogenous covariates (this is an extension of the Berry transformation we introduced last time).
3. Nesting the random utility model within a larger structural model of imperfect competition in order to obtain consistent estimation of underlying structural parameters.

Endogenous attributes

- Often we are working with limited information/data on attributes that consumers care about.
- The standard mixed logit specification assumes all attributes (X_{it}) are independent of unobservables captured in the error term.
- When thinking about the demand for differentiated products, it seems likely that this assumption will be violated... why?

Endogenous attributes

- Often we are working with limited information/data on attributes that consumers care about.
- The standard mixed logit specification assumes all attributes (X_{it}) are independent of unobservables captured in the error term.
- When thinking about the demand for differentiated products, it seems likely that this assumption will be violated... why?
- For example: product attributes that are valued by consumers but are unobserved by the researcher can be correlated with prices (and other observed attributes).

Reality bites: Endogeneity concerns

- **Problem:** Prices and non-price attributes are determined by profit-maximizing firms who have access to information that we (researchers) do not. If unobserved attributes correlated with observed attributes, we've got an endogeneity problem!
- An important insight of Berry (1997) is that we can move problem 1 into the "observed" portion of utility by including product specific constants.

IV strategy?

Note that the product specific constant δ_{jt} capture the average utility associated with both observed *and unobserved* attributes of product j in market t.

$$\delta_{jt} = X_{jt}\beta + \alpha p_{jt} + \xi_{jt}$$

- These capture the average valuation of all characteristics specific to a product.
- Once these product-specific dummies are included, the ε error no longer contains unobserved product characteristics. Rather, the error contains deviations from the unobserved average utility.
- The average valuation of these omitted/unobserved product attributes are now subsumed into the δ_{jt};

IV strategy

$$\delta_{jt} = X_{jt}\beta + \alpha p_{jt} + \xi_{jt}$$

- The δ_{jt} are linear in the structural error ξ_{jt}.
- This allows us to smuggle the endogeneity problem into a linear setting where the endogeneity issue can be addressed using standard methods (IV).
- Once you have estimated your model-specific δ_{jt}, these become the dependent variable in an auxiliary regression.
- Regress the δ_{jt} on vehicle attributes and vehicle prices. In this linear setting, familiar IV strategies can be used.

We need variation that affects pricing of different brands or models differently, but is uncorrelated with demand shocks. Ideas??

With many products, you quickly encounter computational complications

- With a large number of products and markets, we have a large number of δ_{jt} to estimate.
- Within a simple conditional logit framework, the δ_{jt} can be estimated as $lnS_{jt} - lnS_{0t}$.
- For the mixed logit specification, the system of share equations is non-linear and must be solved numerically.
- Estimation of a large number of constants can be numerically difficult using standard gradient-based optimization algorithms.
- BLP provide us with a more tractable way to estimate many mean utility parameters.....

Innovation: Contraction mapping

- The key intuition, building on Berry (1994), is that for any value of θ_2, there exists a unique δ vector such that the predicted shares equal actual (observed) shares.
- Estimation therefore involves finding the δ that matches observed market shares given assumed θ_2.
- This can be hard - you may be trying to estimate zillions of !
- BLP demonstrate an iterative contraction mapping procedure that finds the vector δ that equates the observed and predicted market shares for a given θ_2.

Basic idea behind contraction mapping

1. Pick an initial guess of δ .
2. Conditional on this δ vector and assumed set of θ_2 parameters, compute (via simulation) the estimated shares.
3. Using these predicted shares, they then update the δ vector using:

$$\delta_j^n = \delta_j^{n-1} + \ln\left(\frac{s_j}{\widehat{s}_j(\delta_j^{n-1};\theta_2)}\right).$$

4. Recompute the market shares with this updated δ and repeat until the δ converges to the solution

The empirical solution yields an empirical approximation to the non-linear market share functions.

Berry transformation and the contraction mapping

- The empirical solution to this contraction mapping yields an empirical approximation to the non-linear market share functions.
- This allows us to solve numerically for the δ parameters.
- Any algorithm that can solve numerically for a system of equations can do the trick.
- **Note that this contraction mapping does require accurate estimation of market shares. So market definition matters!**

BLP

In this lecture, we will be focusing on three BLP innovations in particular:

1. Estimating models of consumer-level demand using aggregate (i.e. market level) versus disaggregated (i.e. household or customer level) data (hereby generalizing the use of simulation methods)
2. Demonstrating a tractable means of dealing with endogenous covariates (this is an extension of the Berry transformation we introduced last time).
3. Nesting the random utility model within a larger structural model of imperfect competition in order to obtain consistent estimation of underlying structural parameters.

What about the supply side?

- Thus far, we have emphasized the demand-side of these differentiated product market models.
- BLP (and many subsequent applications) integrate a model of strategic behavior on the supply side.
  - Simulating market outcomes under changed conditions requires a model of supply as well as demand– prices are determined by the interaction of demand *and* supply!
  - n this case, the supply side is also important in our search for valid instruments (we need a theory of how the supply-side is choosing prices)
- Incorporating the supply side has the potential to improve the estimates of demand and expand the usefulness of the model....but requires more structure!

Bertrand competition

- BLP assume firms engage in Bertrand price competition.
- Because products are differentiated, Bertrand competition does not imply marginal cost pricing.
- Sellers know the demand system they face and the constant marginal costs of their rivals. Each firm sets prices so as to maximize profits, given the prices and costs of other firms.
- Key assumption is that product lines (and non-price attributes) are fixed.

Bertrand FOCs

Let p denote the vector of product prices, mc denote the vector of marginal costs, and Ω denote a kxk matrix of $\frac{dq_j}{dp_i}$ elements, $s(p)$ is the vector of market shares .

The first-order conditions for a static Bertrand model with multi-product firms:

$$p - mc = \Omega(p)^{-1}s(p),$$

- The elements of the Ω matrix capture the derivatives of market shares with respect to prices.
- Demand estimation gives us all the elements we need to construct this Ω matrix .
- These FOC can be inverted to recover marginal cost estimates!

Role of the supply side

- By invoking all this structure, BLP can now estimate the primitives for the demand-side *and the supply-side (marginal costs).*
- Using additional structure from the supply-side forces the estimation to find cost AND demand parameters that are consistent with profit maximization.
- Estimated marginal costs and demand parameter estimates are those that rationalize observed market shares conditional on our assumptions about the nature of the strategic interactions between firms, the assumptions about the structure of the utility function, the structure of the marginal cost function.

How does the supply side model help address our endogeneity problem?

- Hint: Build on the Bresnahan (1987) argument that the location of products in attribute space is fixed/determined before the pricing decision.
- In the Bertrand model, firms choose attributes and prices to maximize profits subject to the constraints imposed by their cost structure and the residual demand they face.
- In other words: each manufacturer prices each of its products in a way that takes into account substitution patterns with its other products, and with the products of rival firms.

Ideas?

BLP instruments

Under Bertrand oligopoly pricing, products that face good substitutes will have low markups whereas other products will have high markups. Nash markups will respond differently to own and rival products.

BLP instruments:

- Own-product non-price attributes
- Sum of non-price attributes for other products offered by firm.
- Sum of non-price characteristics for producs offered by rivals.

If we assume non-price attributes (e.g. fuel economy) are exogenous to the model, the fuel economy of vehicle j will be correlated with the price of vehicle i, but uncorrelated with the unobservable attributes of vehicle i that gave rise to the endogeneity problem in the first place.

Concerns??

Concerns??

"It is important not to get carried away by the technical fireworks, and to remember this most basic, yet very difficult, identification problem."

The validity of these instruments depends critically on the assumption that other product characteristics are uncorrelated with the unobserved product attributes.

- Non-price attributes can also be endogenous! Vehicle attributes such as acceleration and fuel economy can be modified in the short term by manufacturers.

More grumbling..

- Truly exogenous non-price attribute assumption is hard to justify.
- With product fixed effects, there may be little variation left in the instruments (exit and entry becomes really (too?) important)
- Another concern.. instruments do not identify the elements of the covariance matrix (we'll come back to this).

Identification recap:

- Linear parameters that enter mean utilities are recovered by regressing estimated mean utility δs on product attributes.
- This is just like the hedonic regression papers we looked at. But now the LHS is a measure of mean utility versus price.
- Variation in choice attributes across markets is thus important and the main source of identifying variation!
- Changes in relative prices/attributes across markets is key.

Estimation overview

- Collect data on product attributes X and market shares y.
- Specify latent utility function. Make your parametric assumptions about how random taste parameters are distributed in the population.
- Simulate drawing from these distributions (generate a set of S draws)
- Estimate the implied market shares given these draws and an initial guess of the mean utilities and an initial guess of the non-linear parameters.
- Inversion step recovers the mean utilities conditional on/implied by these guesses.
- Use mean utilities to form the moment conditions and GMM objective function (or calibrate likelihood function using SMLE)
- Optimize numerically to recover the estimates of non-linear parameters
- Now take your estimated outside and recover the linear parameters given the optimal non-linear parameters.

Returning to that 'mean independence' assumption

BLP assume that the ξ_{jt} are mean independent of the observed attributes of new cars.

Question: If your price coefficient is interacted with demographics and you think prices are endogenous, won't interaction terms also be endogenous ?

Answer: Yes if you think the structural error term is product and demographic specific!! That is, the structural errors are ξ_{cj} (specific to a consumer class).

Solution? To address this, I think you could define separate markets for separate customer types?