Demand for Differentiated Products

Empirical Industrial Organization

Andras Niedermayer[1]

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

Winter 2022

Outline

Theory and estimation with micro data

Theory and estimation with aggregate data

Outline

Theory and estimation with micro data

Theory and estimation with aggregate data

Demand for differentiated products

Products are differentiated

Very few exception: commodities (oil, gold, materials...)

Demand for a product depends on prices of the competing products

Understanding demand for a product includes understanding substitution across products:

- Own price elasticity
- Cross price elasticities

Demand for differentiated products

Products are differentiated

Very few exception: commodities (oil, gold, materials...)

Demand for a product depends on prices of the competing products

Understanding demand for a product includes understanding substitution across products:

- Own price elasticity
- Cross price elasticities

Demand for differentiated products

Products are differentiated

Very few exception: commodities (oil, gold, materials...)

Demand for a product depends on prices of the competing products

Understanding demand for a product includes understanding substitution across products:

- Own price elasticity
- Cross price elasticities

Demand for differentiated products

Products are differentiated

Very few exception: commodities (oil, gold, materials...)

Demand for a product depends on prices of the competing products

Understanding demand for a product includes understanding substitution across products:

- Own price elasticity
- Cross price elasticities

Demand for differentiated products

Products are differentiated

Very few exception: commodities (oil, gold, materials...)

Demand for a product depends on prices of the competing products

Understanding demand for a product includes understanding substitution across products:

- Own price elasticity
- Cross price elasticities

Demand for differentiated products

Products are differentiated

Very few exception: commodities (oil, gold, materials...)

Demand for a product depends on prices of the competing products

Understanding demand for a product includes understanding substitution across products:

- Own price elasticity
- Cross price elasticities

Why do we want to estimate demand?

Understanding demand is crucial for firms

From demand and prices, we can infer margins, mark up which indicate firms' market power

Market power crucial for competition authorities, policy maker etc...

Why do we want to estimate demand?

Understanding demand is crucial for firms

From demand and prices, we can infer margins, mark up which indicate firms' market power

Market power crucial for competition authorities, policy maker etc...

Why do we want to estimate demand?

Understanding demand is crucial for firms

From demand and prices, we can infer margins, mark up which indicate firms' market power

Market power crucial for competition authorities, policy maker etc...

Curse of dimensionality

One could think about estimating:

$$
\begin{aligned}
\ln Q_1 &= \alpha_0^1 + \alpha_1^1 \ln p_1 + \alpha_2^1 \ln p_2 + \cdots + \alpha_N^1 \ln p_N + \epsilon_1 \\
\ln Q_2 &= \alpha_0^2 + \alpha_1^2 \ln p_1 + \alpha_2^2 \ln p_2 + \cdots + \alpha_N^2 \ln p_N + \epsilon_2 \\
&\vdots \\
\ln Q_N &= \alpha_0^N + \alpha_1^N \ln p_1 + \alpha_2^N \ln p_2 + \cdots + \alpha_N^N \ln p_N + \epsilon_N
\end{aligned}
$$

Problem = curse of dimensionality: $(N+1)^2$ parameters to estimate

Model not micro-founded

Curse of dimensionality

One could think about estimating:

$$
\begin{aligned}
\ln Q_1 &= \alpha_0^1 + \alpha_1^1 \ln p_1 + \alpha_2^1 \ln p_2 + \cdots + \alpha_N^1 \ln p_N + \epsilon_1 \\
\ln Q_2 &= \alpha_0^2 + \alpha_1^2 \ln p_1 + \alpha_2^2 \ln p_2 + \cdots + \alpha_N^2 \ln p_N + \epsilon_2 \\
&\vdots \\
\ln Q_N &= \alpha_0^N + \alpha_1^N \ln p_1 + \alpha_2^N \ln p_2 + \cdots + \alpha_N^N \ln p_N + \epsilon_N
\end{aligned}
$$

Problem = curse of dimensionality: $(N+1)^2$ parameters to estimate

Model not micro-founded

Curse of dimensionality

One could think about estimating:

$$
\begin{aligned}
\ln Q_1 &= \alpha_0^1 + \alpha_1^1 \ln p_1 + \alpha_2^1 \ln p_2 + \cdots + \alpha_N^1 \ln p_N + \epsilon_1 \\
\ln Q_2 &= \alpha_0^2 + \alpha_1^2 \ln p_1 + \alpha_2^2 \ln p_2 + \cdots + \alpha_N^2 \ln p_N + \epsilon_2 \\
&\vdots \\
\ln Q_N &= \alpha_0^N + \alpha_1^N \ln p_1 + \alpha_2^N \ln p_2 + \cdots + \alpha_N^N \ln p_N + \epsilon_N
\end{aligned}
$$

Problem = curse of dimensionality: $(N+1)^2$ parameters to estimate

Model not micro-founded

Discrete choice model

Daniel McFadden's multinomial logit

Starts from specification of utility, as a function of characteristics of products (including price) and characteristics of individuals

Idea = move from product space to characteristics space (smaller!)

Each consumer chooses one product among the products available, one option is not to buy

Each consumer chooses the option associated to the highest utility

From the observation of choices and assuming the choice optimality, we can estimate parameters of preferences

Discrete choice model

Daniel McFadden's multinomial logit

Starts from specification of utility, as a function of characteristics of products (including price) and characteristics of individuals

Idea = move from product space to characteristics space (smaller!)

Each consumer chooses one product among the products available, one option is not to buy

Each consumer chooses the option associated to the highest utility

From the observation of choices and assuming the choice optimality, we can estimate parameters of preferences

Discrete choice model

Daniel McFadden's multinomial logit

Starts from specification of utility, as a function of characteristics of products (including price) and characteristics of individuals

Idea = move from product space to characteristics space (smaller!)

Each consumer chooses one product among the products available, one option is not to buy

Each consumer chooses the option associated to the highest utility

From the observation of choices and assuming the choice optimality, we can estimate parameters of preferences

Discrete choice model

Daniel McFadden's multinomial logit

Starts from specification of utility, as a function of characteristics of products (including price) and characteristics of individuals

Idea = move from product space to characteristics space (smaller!)

Each consumer chooses one product among the products available, one option is not to buy

Each consumer chooses the option associated to the highest utility

From the observation of choices and assuming the choice optimality, we can estimate parameters of preferences

Discrete choice model

Daniel McFadden's multinomial logit

Starts from specification of utility, as a function of characteristics of products (including price) and characteristics of individuals

Idea = move from product space to characteristics space (smaller!)

Each consumer chooses one product among the products available, one option is not to buy

Each consumer chooses the option associated to the highest utility

From the observation of choices and assuming the choice optimality, we can estimate parameters of preferences

Discrete choice model

Daniel McFadden's multinomial logit

Starts from specification of utility, as a function of characteristics of products (including price) and characteristics of individuals

Idea = move from product space to characteristics space (smaller!)

Each consumer chooses one product among the products available, one option is not to buy

Each consumer chooses the option associated to the highest utility

From the observation of choices and assuming the choice optimality, we can estimate parameters of preferences

Presentation of the model

Utility of product k:

$$\begin{aligned} U_{ik} &= X_k \beta_i + \varepsilon_{ik} \\ &= \delta_k + \mu_{ik} + \varepsilon_{ik} \end{aligned}$$

where X are product characteristics (e.g size, brand, colour, ingredients, price...)

β_i represents the valuation of the characteristics, it is the vector of parameters of interest

δ_k represents the mean utility of product k (common to all individuals)

μ_{ik} represents individual and product-specific variables

ε_{ik} is the error term, assumed to be iid across products and individuals

Presentation of the model

Utility of product k:

$$\begin{aligned} U_{ik} &= X_k \beta_i + \varepsilon_{ik} \\ &= \delta_k + \mu_{ik} + \varepsilon_{ik} \end{aligned}$$

where X are product characteristics (e.g size, brand, colour, ingredients, price...)

β_i represents the valuation of the characteristics, it is the vector of parameters of interest

δ_k represents the mean utility of product k (common to all individuals)

μ_{ik} represents individual and product-specific variables

ε_{ik} is the error term, assumed to be iid across products and individuals

Presentation of the model

Utility of product k:

$$\begin{aligned} U_{ik} &= X_k \beta_i + \varepsilon_{ik} \\ &= \delta_k + \mu_{ik} + \varepsilon_{ik} \end{aligned}$$

where X are product characteristics (e.g size, brand, colour, ingredients, price...)

β_i represents the valuation of the characteristics, it is the vector of parameters of interest

δ_k represents the mean utility of product k (common to all individuals)

μ_{ik} represents individual and product-specific variables

ε_{ik} is the error term, assumed to be iid across products and individuals

Presentation of the model

Utility of product k:

$$\begin{aligned} U_{ik} &= X_k \beta_i + \varepsilon_{ik} \\ &= \delta_k + \mu_{ik} + \varepsilon_{ik} \end{aligned}$$

where X are product characteristics (e.g size, brand, colour, ingredients, price...)

β_i represents the valuation of the characteristics, it is the vector of parameters of interest

δ_k represents the mean utility of product k (common to all individuals)

μ_{ik} represents individual and product-specific variables

ε_{ik} is the error term, assumed to be iid across products and individuals

Presentation of the model

Utility of product k:

$$\begin{aligned} U_{ik} &= X_k \beta_i + \varepsilon_{ik} \\ &= \delta_k + \mu_{ik} + \varepsilon_{ik} \end{aligned}$$

where X are product characteristics (e.g size, brand, colour, ingredients, price...)

β_i represents the valuation of the characteristics, it is the vector of parameters of interest

δ_k represents the mean utility of product k (common to all individuals)

μ_{ik} represents individual and product-specific variables

ε_{ik} is the error term, assumed to be iid across products and individuals

The logit model

The error terms ε_{ik} are assumed to be extreme value distributed

Cumulative distribution function:

$$F_\varepsilon(x) = \exp[-\exp(-x)]$$

Density function:

$$f_\varepsilon(x) = \exp(-x)\exp[-\exp(-x)]$$

The probability that consumer i chooses product j has a closed-form solution:

$$P_{ij} = \frac{\exp(\delta_j + \mu_{ij})}{\sum_{k=0}^{J}\exp(\delta_k + \mu_{ik})}$$

Derived from utility maximization, if ε_{ik} are iid extreme value distributed

The logit model

The error terms ε_{ik} are assumed to be extreme value distributed

Cumulative distribution function:

$$F_\varepsilon(x) = \exp[-\exp(-x)]$$

Density function:

$$f_\varepsilon(x) = \exp(-x)\exp[-\exp(-x)]$$

The probability that consumer i chooses product j has a closed-form solution:

$$P_{ij} = \frac{\exp(\delta_j + \mu_{ij})}{\sum_{k=0}^{J} \exp(\delta_k + \mu_{ik})}$$

Derived from utility maximization, if ε_{ik} are iid extreme value distributed

The logit model

The error terms ε_{ik} are assumed to be extreme value distributed

Cumulative distribution function:

$$F_\varepsilon(x) = \exp[-\exp(-x)]$$

Density function:

$$f_\varepsilon(x) = \exp(-x)\exp[-\exp(-x)]$$

The probability that consumer i chooses product j has a closed-form solution:

$$P_{ij} = \frac{\exp(\delta_j + \mu_{ij})}{\sum_{k=0}^{J} \exp(\delta_k + \mu_{ik})}$$

Derived from utility maximization, if ε_{ik} are iid extreme value distributed

The logit model

The error terms ε_{ik} are assumed to be extreme value distributed

Cumulative distribution function:

$$F_\varepsilon(x) = \exp[-\exp(-x)]$$

Density function:

$$f_\varepsilon(x) = \exp(-x)\exp[-\exp(-x)]$$

The probability that consumer i chooses product j has a closed-form solution:

$$P_{ij} = \frac{\exp(\delta_j + \mu_{ij})}{\sum_{k=0}^{J} \exp(\delta_k + \mu_{ik})}$$

Derived from utility maximization, if ε_{ik} are iid extreme value distributed

The logit model

The error terms ε_{ik} are assumed to be extreme value distributed

Cumulative distribution function:

$$F_\varepsilon(x) = \exp[-\exp(-x)]$$

Density function:

$$f_\varepsilon(x) = \exp(-x)\exp[-\exp(-x)]$$

The probability that consumer i chooses product j has a closed-form solution:

$$P_{ij} = \frac{\exp(\delta_j + \mu_{ij})}{\sum_{k=0}^{J} \exp(\delta_k + \mu_{ik})}$$

Derived from utility maximization, if ε_{ik} are iid extreme value distributed

Proof in a simple case

Proof with 3 products: $J = 3$, assume $\mu_{ik} = 0$

$$
\begin{aligned}
& \Pr(\text{choice} \; = 2) \\
= \; & \Pr(U_2 > U_1 \text{ and } U_2 > U_3) \\
= \; & \Pr(\varepsilon_1 < \varepsilon_2 + \delta_2 - \delta_1 \text{ and } \varepsilon_3 < \varepsilon_2 + \delta_2 - \delta_3) \\
= \; & \int_{-\infty}^{+\infty} f_\varepsilon(x) \Pr(\varepsilon_1 < \varepsilon_2 + \delta_2 - \delta_1 \text{ and } \varepsilon_3 < \varepsilon_2 + \delta_2 - \delta_3 | \boldsymbol{\varepsilon_2} = \mathbf{x}) dx \\
= \; & \int_{-\infty}^{\infty} f_\varepsilon(x) \left[\int_{-\infty}^{x+\delta_2-\delta_1} f(\varepsilon_1) d\varepsilon_1 \cdot \int_{-\infty}^{x+\delta_2-\delta_3} f(\varepsilon_3) d\varepsilon_3 \right] dx \\
= \; & \int_{-\infty}^{\infty} \exp(-x) \exp[-\exp(-x)] \cdot F_\varepsilon(x + \delta_2 - \delta_1) \cdot F_\varepsilon(x + \delta_2 - \delta_3) dx \\
= \; & \int_{-\infty}^{\infty} \exp(-x) \exp[-\exp(-x)] \cdot \exp[-\exp(-x - \delta_2 + \delta_1)] \\
& \cdot \exp[-\exp(-x - \delta_2 + \delta_3)] dx
\end{aligned}
$$

Proof in a simple case

Proof cont'd

$$= \int_{-\infty}^{\infty} \exp(-x) \exp\left(-\exp(-x).\left[1+\exp(\delta_1-\delta_2)+\exp(\delta_3-\delta_2)\right]\right) dx$$

Change of variable: $t = \exp(-x)$, $dt = -\exp(-x)dx$

$$
\begin{aligned}
&= \int_{+\infty}^{0} -dt \exp\left(-t\left[1+\exp(\delta_1-\delta_2)+\exp(\delta_3-\delta_2)\right]\right) \\
&= \int_{0}^{+\infty} \exp\left(-t\left[1+\exp(\delta_1-\delta_2)+\exp(\delta_3-\delta_2)\right]\right) dt \\
&= \left[\frac{\exp\left(-t\left[1+\exp(\delta_1-\delta_2)+\exp(\delta_3-\delta_2)\right]\right)}{-(1+\exp(\delta_1-\delta_2)+\exp(\delta_3-\delta_2))}\right]_0^{+\infty} \\
&= 0 - \frac{1}{-(1+\exp(\delta_1-\delta_2)+\exp(\delta_3-\delta_2))} \\
&= \frac{\exp(\delta_2)}{\exp(\delta_1)+\exp(\delta_2)+\exp(\delta_3)}
\end{aligned}
$$

Proof in a simple case

Proof cont'd

$$= \int_{-\infty}^{\infty} \exp(-x) \exp\left(-\exp(-x).\left[1+\exp(\delta_1-\delta_2)+\exp(\delta_3-\delta_2)\right]\right) dx$$

Change of variable: $t=\exp(-x)$, $dt=-\exp(-x)dx$

$$\begin{aligned}
&= \int_{+\infty}^{0} -dt \exp\left(-t\left[1+\exp(\delta_1-\delta_2)+\exp(\delta_3-\delta_2)\right]\right) \\
&= \int_{0}^{+\infty} \exp\left(-t\left[1+\exp(\delta_1-\delta_2)+\exp(\delta_3-\delta_2)\right]\right) dt \\
&= \left[\frac{\exp\left(-t\left[1+\exp(\delta_1-\delta_2)+\exp(\delta_3-\delta_2)\right]\right)}{-(1+\exp(\delta_1-\delta_2)+\exp(\delta_3-\delta_2))}\right]_0^{+\infty} \\
&= 0-\frac{1}{-(1+\exp(\delta_1-\delta_2)+\exp(\delta_3-\delta_2))} \\
&= \frac{\exp(\delta_2)}{\exp(\delta_1)+\exp(\delta_2)+\exp(\delta_3)}
\end{aligned}$$

Estimation of the model

Assume we observe choices of N individuals and the choice set

$D_{ik} = 1$ if product k is purchased by individual i, $D_{ik} = 0$ if it is not chosen

Estimation by maximum likelihood

$\max_{\beta} LL = \log \left(\prod_i \prod_j \left(P_{ij}(X, \beta) \right)^{D_{ij}} \right)$

Challenges:

- No purchase typically not observed, solutions:
  - model the choice of a product, conditional on buying
  - make an assumption on number (and characteristics) of consumers who did not buy any product (= define market potential)
- Consistency requires exogeneity of X, not true for the price
- Unobserved product characteristics? Can be captured with a product dummy (need large number of consumers relative to the number of products)

Estimation of the model

Assume we observe choices of N individuals and the choice set $D_{ik} = 1$ if product k is purchased by individual i, $D_{ik} = 0$ if it is not chosen

Estimation by maximum likelihood

$\max_\beta LL = \log\left(\prod_i \prod_j \left(P_{ij}(X,\beta)\right)^{D_{ij}}\right)$

Challenges:

- No purchase typically not observed, solutions:
  - model the choice of a product, conditional on buying
  - make an assumption on number (and characteristics) of consumers who did not buy any product (=define market potential)
- Consistency requires exogeneity of X, not true for the price
- Unobserved product characteristics? Can be captured with a product dummy (need large number of consumers relative to the number of products)

Estimation of the model

Assume we observe choices of N individuals and the choice set $D_{ik} = 1$ if product k is purchased by individual i, $D_{ik} = 0$ if it is not chosen

Estimation by maximum likelihood

$$\max_{\beta} LL = \log \left(\prod_i \prod_j \left(P_{ij}(X, \beta) \right)^{D_{ij}} \right)$$

Challenges:

- No purchase typically not observed, solutions:
  - model the choice of a product, conditional on buying
  - make an assumption on number (and characteristics) of consumers who did not buy any product – define market potential)
- Consistency requires exogeneity of X, not true for the price
- Unobserved product characteristics? Can be captured with a product dummy (need large number of consumers relative to the number of products)

Estimation of the model

Assume we observe choices of N individuals and the choice set $D_{ik} = 1$ if product k is purchased by individual i, $D_{ik} = 0$ if it is not chosen

Estimation by maximum likelihood

$\max_{\beta} LL = \log \left(\prod_i \prod_j \left(P_{ij}(X, \beta) \right)^{D_{ij}} \right)$

Challenges:

- No purchase typically not observed, solutions:
  - model the choice of a product, conditional on buying
  - make an assumption on number (and characteristics) of consumers who did not buy any product (= define a market potential)
- Consistency requires exogeneity of X, not true for the price
- Unobserved product characteristics? Can be captured with a product dummy (need large number of consumers relative to the number of products)

Estimation of the model

Assume we observe choices of N individuals and the choice set $D_{ik} = 1$ if product k is purchased by individual i, $D_{ik} = 0$ if it is not chosen

Estimation by maximum likelihood

$\max_\beta LL = \log \left(\prod_i \prod_j \left(P_{ij}(X, \beta) \right)^{D_{ij}} \right)$

Challenges:

- No purchase typically not observed, solutions:
  - model the choice of a product, conditional on buying
  - make an assumption on number (and characteristics) of consumers who did not buy any product (= define a market potential)
- Consistency requires exogeneity of X, not true for the price
- Unobserved product characteristics? Can be captured with a product dummy (need large number of consumers relative to the number of products)

Estimation of the model

Assume we observe choices of N individuals and the choice set $D_{ik} = 1$ if product k is purchased by individual i, $D_{ik} = 0$ if it is not chosen

Estimation by maximum likelihood

$\max_\beta LL = \log \left(\prod_i \prod_j \left(P_{ij}(X, \beta)\right)^{D_{ij}} \right)$

Challenges:

- No purchase typically not observed, solutions:
  - model the choice of a product, conditional on buying
  - make an assumption on number (and characteristics) of consumers who did not buy any product (= define a market potential)
- Consistency requires exogeneity of X, not true for the price
- Unobserved product characteristics? Can be captured with a product dummy (need large number of consumers relative to the number of products)

Estimation of the model

Assume we observe choices of N individuals and the choice set $D_{ik} = 1$ if product k is purchased by individual i, $D_{ik} = 0$ if it is not chosen

Estimation by maximum likelihood

$\max_\beta LL = \log \left(\prod_i \prod_j \left(P_{ij}(X, \beta) \right)^{D_{ij}} \right)$

Challenges:

- No purchase typically not observed, solutions:
  - model the choice of a product, conditional on buying
  - make an assumption on number (and characteristics) of consumers who did not buy any product (= define a market potential)
- Consistency requires exogeneity of X, not true for the price
- Unobserved product characteristics? Can be captured with a product dummy (need large number of consumers relative to the number of products)

Estimation of the model

Assume we observe choices of N individuals and the choice set $D_{ik} = 1$ if product k is purchased by individual i, $D_{ik} = 0$ if it is not chosen

Estimation by maximum likelihood

$\max_\beta LL = \log \left(\prod_i \prod_j \left(P_{ij}(X, \beta) \right)^{D_{ij}} \right)$

Challenges:

- No purchase typically not observed, solutions:
  - model the choice of a product, conditional on buying
  - make an assumption on number (and characteristics) of consumers who did not buy any product (= define a market potential)
- Consistency requires exogeneity of X, not true for the price
- Unobserved product characteristics? Can be captured with a product dummy (need large number of consumers relative to the number of products)

Estimation of the model

What is μ_{ij}?

Depends of the data:

- Repeated purchases (panel of individual choices in different environments): we can identify μ_{ij} with product $\times$ individual fixed effect
- Repeated cross-section of individual choices: we need to specify μ_{ij}

$$\mu_{ij} = D_i \pi X_j$$

  with D_i some observable individual characteristics (age, income...) and π some heterogeneity parameters to estimate

Estimation of the model

What is μ_{ij}?

Depends of the data:

- Repeated purchases (panel of individual choices in different environments): we can identify μ_{ij} with product $\times$ individual fixed effect
- Repeated cross-section of individual choices: we need to specify μ_{ij}

$$\mu_{ij} = D_i \pi X_j$$

  with D_i some observable individual characteristics (age, income...) and π some heterogeneity parameters to estimate

Estimation of the model

What is μ_{ij}?

Depends of the data:

- Repeated purchases (panel of individual choices in different environments): we can identify μ_{ij} with product $\times$ individual fixed effect
- Repeated cross-section of individual choices: we need to specify μ_{ij}

$$\mu_{ij} = D_i \pi X_j$$

  with D_i some observable individual characteristics (age, income...) and π some heterogeneity parameters to estimate

Estimation of the model

What is μ_{ij}?

Depends of the data:

- Repeated purchases (panel of individual choices in different environments): we can identify μ_{ij} with product $\times$ individual fixed effect
- Repeated cross-section of individual choices: we need to specify μ_{ij}

  $$\mu_{ij} = D_i \pi X_j$$

  with D_i some observable individual characteristics (age, income...) and π some heterogeneity parameters to estimate

Random coefficients

What if we do not observe individual characteristics?

We can introduce **unobserved heterogeneity** through random coefficients

We specify the distribution of β and estimate its parameters

Example, we specify β as normally distributed so we have to estimate the mean and standard deviation $\bar{\beta}$ and σ^{β}

We can always re-write draws from $N(\bar{\beta}, \sigma^{\beta})$ as:

$$\beta_s = \bar{\beta} + \sigma^{\beta}\nu_s$$

Idea: from observed purchases, back out how heterogeneous are preferences for certain product attributes, identification leverages repeated choices under different choice sets

Random coefficients

What if we do not observe individual characteristics?

We can introduce **unobserved heterogeneity** through random coefficients

We specify the distribution of β and estimate its parameters

Example, we specify β as normally distributed so we have to estimate the mean and standard deviation $\bar{\beta}$ and σ^{β}

We can always re-write draws from $N(\bar{\beta}, \sigma^{\beta})$ as:

$$\beta_s = \bar{\beta} + \sigma^{\beta} \nu_s$$

Idea: from observed purchases, back out how heterogeneous are preferences for certain product attributes, identification leverages repeated choices under different choice sets

Random coefficients

What if we do not observe individual characteristics?

We can introduce **unobserved heterogeneity** through random coefficients

We specify the distribution of β and estimate its parameters

Example, we specify β as normally distributed so we have to estimate the mean and standard deviation $\bar{\beta}$ and σ^β

We can always re-write draws from $N(\bar{\beta}, \sigma^\beta)$ as:

$$\beta_s = \bar{\beta} + \sigma^\beta \nu_s$$

Idea: from observed purchases, back out how heterogeneous are preferences for certain product attributes, identification leverages repeated choices under different choice sets

Random coefficients

What if we do not observe individual characteristics?

We can introduce **unobserved heterogeneity** through random coefficients

We specify the distribution of β and estimate its parameters

Example, we specify β as normally distributed so we have to estimate the mean and standard deviation $\bar{\beta}$ and σ^{β}

We can always re-write draws from $N(\bar{\beta}, \sigma^{\beta})$ as:

$$\beta_s = \bar{\beta} + \sigma^{\beta} \nu_s$$

Idea: from observed purchases, back out how heterogeneous are preferences for certain product attributes, identification leverages repeated choices under different choice sets

Random coefficients

What if we do not observe individual characteristics?

We can introduce **unobserved heterogeneity** through random coefficients

We specify the distribution of β and estimate its parameters

Example, we specify β as normally distributed so we have to estimate the mean and standard deviation $\bar{\beta}$ and σ^{β}

We can always re-write draws from $N(\bar{\beta}, \sigma^{\beta})$ as:

$$\beta_s = \bar{\beta} + \sigma^{\beta} \nu_s$$

Idea: from observed purchases, back out how heterogeneous are preferences for certain product attributes, identification leverages repeated choices under different choice sets

Random coefficients

What if we do not observe individual characteristics?

We can introduce **unobserved heterogeneity** through random coefficients

We specify the distribution of β and estimate its parameters

Example, we specify β as normally distributed so we have to estimate the mean and standard deviation $\bar{\beta}$ and σ^{β}

We can always re-write draws from $N(\bar{\beta}, \sigma^{\beta})$ as:

$$\beta_s = \bar{\beta} + \sigma^{\beta} \nu_s$$

Idea: from observed purchases, back out how heterogeneous are preferences for certain product attributes, identification leverages repeated choices under different choice sets

Random coefficients

The probability that consumer i chooses j has no longer closed-form solution:

$$P_{ij} = \int \frac{\exp(\delta_j + \mu_{ij}(\nu))}{\sum_{k=0}^{J} \exp(\delta_k + \mu_{ik}(\nu))} dF(\nu)$$

And it is the same for every consumer $P_{ij} = P_j \; \forall i$

Integral can no longer be analytically computed but rather approximated using simulations:

$$P_{ij} \simeq \frac{1}{ns} \sum_{s=1}^{ns} \frac{\exp(\delta_j + \mu_j(\nu_s))}{\sum_{k=0}^{J} \exp(\delta_k + \mu_k(\nu_s))}$$

Estimation: Simulated Maximum Likelihood (because we compute probabilities through simulations)

Random coefficients

The probability that consumer i chooses j has no longer closed-form solution:

$$P_{ij} = \int \frac{\exp(\delta_j + \mu_{ij}(\nu))}{\sum_{k=0}^{J} \exp(\delta_k + \mu_{ik}(\nu))} dF(\nu)$$

And it is the same for every consumer $P_{ij} = P_j \ \forall i$

Integral can no longer be analytically computed but rather approximated using simulations:

$$P_{ij} \simeq \frac{1}{ns} \sum_{s=1}^{ns} \frac{\exp(\delta_j + \mu_j(\nu_s))}{\sum_{k=0}^{J} \exp(\delta_k + \mu_k(\nu_s))}$$

Estimation: Simulated Maximum Likelihood (because we compute probabilities through simulations)

Random coefficients

The probability that consumer i chooses j has no longer closed-form solution:

$$P_{ij} = \int \frac{\exp(\delta_j + \mu_{ij}(\nu))}{\sum_{k=0}^{J} \exp(\delta_k + \mu_{ik}(\nu))} dF(\nu)$$

And it is the same for every consumer $P_{ij} = P_j \; \forall i$

Integral can no longer be analytically computed but rather approximated using simulations:

$$P_{ij} \simeq \frac{1}{ns} \sum_{s=1}^{ns} \frac{\exp(\delta_j + \mu_j(\nu_s))}{\sum_{k=0}^{J} \exp(\delta_k + \mu_k(\nu_s))}$$

Estimation: Simulated Maximum Likelihood (because we compute probabilities through simulations)

Random coefficients

The probability that consumer i chooses j has no longer closed-form solution:

$$P_{ij} = \int \frac{\exp(\delta_j + \mu_{ij}(\nu))}{\sum_{k=0}^{J} \exp(\delta_k + \mu_{ik}(\nu))} dF(\nu)$$

And it is the same for every consumer $P_{ij} = P_j \; \forall i$

Integral can no longer be analytically computed but rather approximated using simulations:

$$P_{ij} \simeq \frac{1}{ns} \sum_{s=1}^{ns} \frac{\exp(\delta_j + \mu_j(\nu_s))}{\sum_{k=0}^{J} \exp(\delta_k + \mu_k(\nu_s))}$$

Estimation: Simulated Maximum Likelihood (because we compute probabilities through simulations)

Endogeneity

Introduce unobserved characteristics ξ_{ik}:

$$\begin{aligned} U_{ik} &= X_k\beta_i + \xi_{ik} + \varepsilon_{ik} \\ &= \delta_k + \varepsilon_{ik} \end{aligned}$$

Unobserved characteristics likely to be correlated to some X, e.g. the **price**

- Endogeneity since $E(\xi|p) \neq 0$
- Conditional on p, residual is no longer centered (e.g. we expect higher residuals for higher prices)

Can still use max-likelihood method but proceed as follows:

- First estimate δ with product fixed-effects
- Regress δ on X using IV approach
- Drawback: parameter space may become large!

Endogeneity

Introduce unobserved characteristics ξ_{ik}:

$$\begin{aligned} U_{ik} &= X_k \beta_i + \xi_{ik} + \varepsilon_{ik} \\ &= \delta_k + \varepsilon_{ik} \end{aligned}$$

Unobserved characteristics likely to be correlated to some X, e.g. the **price**

- Endogeneity since $E(\xi|p) \neq 0$
- Conditional on p, residual is no longer centered (e.g. we expect higher residuals for higher prices)

Can still use max-likelihood method but proceed as follows:

- First estimate δ with product fixed-effects
- Regress δ on X using IV approach
- Drawback: parameter space may become large!

Endogeneity

Introduce unobserved characteristics ξ_{ik}:

$$\begin{aligned} U_{ik} &= X_k\beta_i + \xi_{ik} + \varepsilon_{ik} \\ &= \delta_k + \varepsilon_{ik} \end{aligned}$$

Unobserved characteristics likely to be correlated to some X, e.g. the **price**

- Endogeneity since $E(\xi|p) \neq 0$
- Conditional on p, residual is no longer centered (e.g. we expect higher residuals for higher prices)

Can still use max-likelihood method but proceed as follows:

- First estimate δ with product fixed-effects
- Regress δ on X using IV approach
- Drawback: parameter space may become large!

Endogeneity

Introduce unobserved characteristics ξ_{ik}:

$$\begin{aligned} U_{ik} &= X_k\beta_i + \xi_{ik} + \varepsilon_{ik} \\ &= \delta_k + \varepsilon_{ik} \end{aligned}$$

Unobserved characteristics likely to be correlated to some X, e.g. the **price**

- Endogeneity since $E(\xi|p) \neq 0$
- Conditional on p, residual is no longer centered (e.g. we expect higher residuals for higher prices)

Can still use max-likelihood method but proceed as follows:

- First estimate δ with product fixed-effects
- Regress δ on X using IV approach
- Drawback: parameter space may become large!

Endogeneity

Introduce unobserved characteristics ξ_{ik}:

$$\begin{aligned} U_{ik} &= X_k\beta_i + \xi_{ik} + \varepsilon_{ik} \\ &= \delta_k + \varepsilon_{ik} \end{aligned}$$

Unobserved characteristics likely to be correlated to some X, e.g. the **price**

- Endogeneity since $E(\xi|p) \neq 0$
- Conditional on p, residual is no longer centered (e.g. we expect higher residuals for higher prices)

Can still use max-likelihood method but proceed as follows:

- First estimate δ with product fixed-effects
- Regress δ on X using IV approach
- Drawback: parameter space may become large!

Endogeneity

Introduce unobserved characteristics ξ_{ik}:

$$\begin{aligned} U_{ik} &= X_k \beta_i + \xi_{ik} + \varepsilon_{ik} \\ &= \delta_k + \varepsilon_{ik} \end{aligned}$$

Unobserved characteristics likely to be correlated to some X, e.g. the **price**

- Endogeneity since $E(\xi | p) \neq 0$
- Conditional on p, residual is no longer centered (e.g. we expect higher residuals for higher prices)

Can still use max-likelihood method but proceed as follows:

- First estimate δ with product fixed-effects
- Regress δ on X using IV approach
- Drawback: parameter space may become large!

Endogeneity

Introduce unobserved characteristics ξ_{ik}:

$$\begin{aligned} U_{ik} &= X_k \beta_i + \xi_{ik} + \varepsilon_{ik} \\ &= \delta_k + \varepsilon_{ik} \end{aligned}$$

Unobserved characteristics likely to be correlated to some X, e.g. the **price**

- Endogeneity since $E(\xi|p) \neq 0$
- Conditional on p, residual is no longer centered (e.g. we expect higher residuals for higher prices)

Can still use max-likelihood method but proceed as follows:

- First estimate δ with product fixed-effects
- Regress δ on X using IV approach
- Drawback: parameter space may become large!

Endogeneity

Alternative: use control function approach

Idea: regress endogeneous variable(s) on instruments, and use the residuals of the regression as an extra explanatory variable: the **control function**

Issues: (i) rely on a specification of the control function
(ii) rely on an assumption of the distribution of errors in the regression and joint distribution of errors of the model and regression

Formalization:

$$U_{ij} = X_j\beta_i + \alpha_i p_{ij} + \xi_{ij} + \varepsilon_{ij}$$

with $\mathbb{E}(\xi_{ij}|p_{ij}) \neq 0$

Assume we have Z_{ij} such that:

$$p_{ij} = f(Z_{ij}, \gamma) + \omega_{ij}$$

with $\mathbb{E}(\omega_{ij}, \xi_{ij}|Z_{ij}) = 0$ and $\mathbb{E}(\omega_{ij}|\xi_{ij}) \neq 0$

Endogeneity

Alternative: use control function approach

Idea: regress endogeneous variable(s) on instruments, and use the residuals of the regression as an extra explanatory variable: the **control function**

Issues: (i) rely on a specification of the control function

(ii) rely on an assumption of the distribution of errors in the regression and joint distribution of errors of the model and regression

Formalization:

$$U_{ij} = X_j \beta_i + \alpha_i p_{ij} + \xi_{ij} + \varepsilon_{ij}$$

with $\mathbb{E}(\xi_{ij} | p_{ij}) \neq 0$

Assume we have Z_{ij} such that:

$$p_{ij} = f(Z_{ij}, \gamma) + \omega_{ij}$$

with $\mathbb{E}(\omega_{ij}, \xi_{ij} | Z_{ij}) = 0$ and $\mathbb{E}(\omega_{ij} | \xi_{ij}) \neq 0$

Endogeneity

Alternative: use control function approach
Idea: regress endogeneous variable(s) on instruments, and use the residuals of the regression as an extra explanatory variable: the **control function**
Issues: (i) rely on a specification of the control function
(ii) rely on an assumption of the distribution of errors in the regression and joint distribution of errors of the model and regression
Formalization:

$$U_{ij} = X_j\beta_i + \alpha_i p_{ij} + \xi_{ij} + \varepsilon_{ij}$$

with $\mathbb{E}(\xi_{ij}|p_{ij}) \neq 0$
Assume we have Z_{ij} such that:

$$p_{ij} = f(Z_{ij}, \gamma) + \omega_{ij}$$

with $\mathbb{E}(\omega_{ij}, \xi_{ij}|Z_{ij}) = 0$ and $\mathbb{E}(\omega_{ij}|\xi_{ij}) \neq 0$

Endogeneity

Alternative: use control function approach
Idea: regress endogeneous variable(s) on instruments, and use the residuals of the regression as an extra explanatory variable: the **control function**
Issues: (i) rely on a specification of the control function
(ii) rely on an assumption of the distribution of errors in the regression and joint distribution of errors of the model and regression
Formalization:

$$U_{ij} = X_j \beta_i + \alpha_i p_{ij} + \xi_{ij} + \varepsilon_{ij}$$

with $\mathbb{E}(\xi_{ij} | p_{ij}) \neq 0$
Assume we have Z_{ij} such that:

$$p_{ij} = f(Z_{ij}, \gamma) + \omega_{ij}$$

with $\mathbb{E}(\omega_{ij}, \xi_{ij} | Z_{ij}) = 0$ and $\mathbb{E}(\omega_{ij} | \xi_{ij}) \neq 0$

Endogeneity

Alternative: use control function approach
Idea: regress endogeneous variable(s) on instruments, and use the residuals of the regression as an extra explanatory variable: the **control function**
Issues: (i) rely on a specification of the control function
(ii) rely on an assumption of the distribution of errors in the regression and joint distribution of errors of the model and regression
Formalization:

$$U_{ij} = X_j \beta_i + \alpha_i p_{ij} + \xi_{ij} + \varepsilon_{ij}$$

with $\mathbb{E}(\xi_{ij} | p_{ij}) \neq 0$
Assume we have Z_{ij} such that:

$$p_{ij} = f(Z_{ij}, \gamma) + \omega_{ij}$$

with $\mathbb{E}(\omega_{ij}, \xi_{ij} | Z_{ij}) = 0$ and $\mathbb{E}(\omega_{ij} | \xi_{ij}) \neq 0$

Control function approach

We define:

$$\xi_{ij} = \underbrace{\mathbb{E}(\xi_{ij}|\omega_{ij})}_{\text{control function}} + \tilde{\xi}_{ij}$$

with $\mathbb{E}(\tilde{\xi}_{ij}|\omega_{ij}) = 0$ so that $\mathbb{E}(\tilde{\xi}_{ij}|p_{ij}) = 0$

We must specify the control function, simplest:

$$CF(\omega_{ij}, \lambda) = \lambda\omega_{ij}$$

Go back to utility:

$$U_{ij} = X_j\beta_i + \alpha_i p_{ij} + \lambda\omega_{ij} + \underbrace{\tilde{\xi}_{ij} + \varepsilon_{ij}}_{\tilde{\varepsilon}_{ij}}$$

with $\mathbb{E}(\tilde{\xi}_{ij}|p_{ij}, \omega_{ij}) = 0$

What is ω_{ij}? Unobserved but replaced by an estimate: $\hat{\omega}_{ij}$ is the residual of the regression of p_{ij} on Z_{ij}

Control function approach

We define:

$$\xi_{ij} = \underbrace{\mathbb{E}(\xi_{ij}|\omega_{ij})}_{\text{control function}} + \tilde{\xi}_{ij}$$

with $\mathbb{E}(\tilde{\xi}_{ij}|\omega_{ij}) = 0$ so that $\mathbb{E}(\tilde{\xi}_{ij}|p_{ij}) = 0$

We must specify the control function, simplest:

$$CF(\omega_{ij}, \lambda) = \lambda\omega_{ij}$$

Go back to utility:

$$U_{ij} = X_j\beta_i + \alpha_i p_{ij} + \lambda\omega_{ij} + \underbrace{\tilde{\xi}_{ij} + \varepsilon_{ij}}_{\tilde{\varepsilon}_{ij}}$$

with $\mathbb{E}(\tilde{\xi}_{ij}|p_{ij}, \omega_{ij}) = 0$

What is ω_{ij}? Unobserved but replaced by an estimate: $\hat{\omega}_{ij}$ is the residual of the regression of p_{ij} on Z_{ij}

Control function approach

We define:

$$\xi_{ij} = \underbrace{\mathbb{E}(\xi_{ij}|\omega_{ij})}_{\text{control function}} + \tilde{\xi}_{ij}$$

with $\mathbb{E}(\tilde{\xi}_{ij}|\omega_{ij}) = 0$ so that $\mathbb{E}(\tilde{\xi}_{ij}|p_{ij}) = 0$

We must specify the control function, simplest:

$$CF(\omega_{ij}, \lambda) = \lambda\omega_{ij}$$

Go back to utility:

$$U_{ij} = X_j\beta_i + \alpha_i p_{ij} + \lambda\omega_{ij} + \underbrace{\tilde{\xi}_{ij} + \varepsilon_{ij}}_{\tilde{\varepsilon}_{ij}}$$

with $\mathbb{E}(\tilde{\xi}_{ij}|p_{ij}, \omega_{ij}) = 0$

What is ω_{ij}? Unobserved but replaced by an estimate: $\hat{\omega}_{ij}$ is the residual of the regression of p_{ij} on Z_{ij}

Control function approach

Probabilities computed as before except we need to integrate on the distribution of $\tilde{\varepsilon}$ conditional on ω

Conditional distribution of the new error term $f(\tilde{\varepsilon}_{ij}|\omega_{ij})$? Depends on the assumption made on the distribution of $\tilde{\xi}$

Assume $\tilde{\xi}$ normally distributed and ε EV distributed, we have a mixed logit, no analytic formula, need to use numerical approximation of the integral through simulations

Reference: Petrin & Train (2009)

Control function approach

Probabilities computed as before except we need to integrate on the distribution of $\tilde{\varepsilon}$ conditional on ω

Conditional distribution of the new error term $f(\tilde{\varepsilon}_{ij}|\omega_{ij})$?
Depends on the assumption made on the distribution of $\tilde{\xi}$

Assume $\tilde{\xi}$ normally distributed and ε EV distributed, we have a mixed logit, no analytic formula, need to use numerical approximation of the integral through simulations

Reference: Petrin & Train (2009)

Control function approach

Probabilities computed as before except we need to integrate on the distribution of $\tilde{\varepsilon}$ conditional on ω

Conditional distribution of the new error term $f(\tilde{\varepsilon}_{ij}|\omega_{ij})$?
Depends on the assumption made on the distribution of $\tilde{\xi}$

Assume $\tilde{\xi}$ normally distributed and ε EV distributed, we have a mixed logit, no analytic formula, need to use numerical approximation of the integral through simulations

Reference: Petrin & Train (2009)

Control function approach

Probabilities computed as before except we need to integrate on the distribution of $\tilde{\varepsilon}$ conditional on ω

Conditional distribution of the new error term $f(\tilde{\varepsilon}_{ij}|\omega_{ij})$?
Depends on the assumption made on the distribution of $\tilde{\xi}$

Assume $\tilde{\xi}$ normally distributed and ε EV distributed, we have a mixed logit, no analytic formula, need to use numerical approximation of the integral through simulations

Reference: Petrin & Train (2009)

Presentation

We do not always have access to consumer-level data

We can estimate demand with aggregate data, i.e. total quantity purchased at the product level

Model from Berry, Levinsohn & Pakes (1995), and Berry (1994)

With product-level aggregate data, using the BLP model, we will still be able to:

- include unobserved product characteristics
- control for price endogeneity
- include unobserved individual heterogeneity

Very influential model, used for research and practitioners (e.g. by competition authority)

Presentation

We do not always have access to consumer-level data

We can estimate demand with aggregate data, i.e. total quantity purchased at the product level

Model from Berry, Levinsohn & Pakes (1995), and Berry (1994)

With product-level aggregate data, using the BLP model, we will still be able to:

- include unobserved product characteristics
- control for price endogeneity
- include unobserved individual heterogeneity

Very influential model, used for research and practitioners (e.g. by competition authority)

Presentation

We do not always have access to consumer-level data

We can estimate demand with aggregate data, i.e. total quantity purchased at the product level

Model from Berry, Levinsohn & Pakes (1995), and Berry (1994)

With product-level aggregate data, using the BLP model, we will still be able to:

- include unobserved product characteristics
- control for price endogeneity
- include unobserved individual heterogeneity

Very influential model, used for research and practitioners (e.g. by competition authority)

Presentation

We do not always have access to consumer-level data

We can estimate demand with aggregate data, i.e. total quantity purchased at the product level

Model from Berry, Levinsohn & Pakes (1995), and Berry (1994)

With product-level aggregate data, using the BLP model, we will still be able to:

- include unobserved product characteristics
- control for price endogeneity
- include unobserved individual heterogeneity

Very influential model, used for research and practitioners (e.g. by competition authority)

Presentation

We do not always have access to consumer-level data

We can estimate demand with aggregate data, i.e. total quantity purchased at the product level

Model from Berry, Levinsohn & Pakes (1995), and Berry (1994)

With product-level aggregate data, using the BLP model, we will still be able to:

- include unobserved product characteristics
- control for price endogeneity
- include unobserved individual heterogeneity

Very influential model, used for research and practitioners (e.g. by competition authority)

Presentation

We do not always have access to consumer-level data

We can estimate demand with aggregate data, i.e. total quantity purchased at the product level

Model from Berry, Levinsohn & Pakes (1995), and Berry (1994)

With product-level aggregate data, using the BLP model, we will still be able to:

- include unobserved product characteristics
- control for price endogeneity
- include unobserved individual heterogeneity

Very influential model, used for research and practitioners (e.g. by competition authority)

Presentation

We do not always have access to consumer-level data

We can estimate demand with aggregate data, i.e. total quantity purchased at the product level

Model from Berry, Levinsohn & Pakes (1995), and Berry (1994)

With product-level aggregate data, using the BLP model, we will still be able to:

- include unobserved product characteristics
- control for price endogeneity
- include unobserved individual heterogeneity

Very influential model, used for research and practitioners (e.g. by competition authority)

Presentation

We do not always have access to consumer-level data

We can estimate demand with aggregate data, i.e. total quantity purchased at the product level

Model from Berry, Levinsohn & Pakes (1995), and Berry (1994)

With product-level aggregate data, using the BLP model, we will still be able to:

- include unobserved product characteristics
- control for price endogeneity
- include unobserved individual heterogeneity

Very influential model, used for research and practitioners (e.g. by competition authority)

The BLP model

$$U_{ik} = X_k\beta_i - \alpha_i p_k + \xi_k + \varepsilon_{ik}$$
$$= \delta_k + \mu_{ik} + \varepsilon_{ik}$$

- X_k: observed product characteristics
- p_k: price of the product, endogenous: $\mathbb{E}(\xi|p) \neq 0$
- ξ_k: unobserved product characteristics (unobserved to the econometrician)
- δ_k: mean product utility
- μ_{ik}: individual deviation from the mean utility
- Important assumption: ξ_k are the same for all consumers, no unobserved heterogeneity in the valuation of the unobserved characteristics
- We examine first two restricted versions of the BLP model: logit and nested logit models

The BLP model

$$U_{ik} = X_k\beta_i - \alpha_i p_k + \xi_k + \varepsilon_{ik}$$
$$= \delta_k + \mu_{ik} + \varepsilon_{ik}$$

- X_k: observed product characteristics
- p_k: price of the product, endogenous: $\mathbb{E}(\xi|p) \neq 0$
- ξ_k: unobserved product characteristics (unobserved to the econometrician)
- δ_k: mean product utility
- μ_{ik}: individual deviation from the mean utility
- Important assumption: ξ_k are the same for all consumers, no unobserved heterogeneity in the valuation of the unobserved characteristics
- We examine first two restricted versions of the BLP model: logit and nested logit models

The BLP model

$$U_{ik} = X_k\beta_i - \alpha_i p_k + \xi_k + \varepsilon_{ik}$$
$$= \delta_k + \mu_{ik} + \varepsilon_{ik}$$

- X_k: observed product characteristics
- p_k: price of the product, endogenous: $\mathbb{E}(\xi|p) \neq 0$
- ξ_k: unobserved product characteristics (unobserved to the econometrician)
- δ_k: mean product utility
- μ_{ik}: individual deviation from the mean utility
- Important assumption: ξ_k are the same for all consumers, no unobserved heterogeneity in the valuation of the unobserved characteristics
- We examine first two restricted versions of the BLP model: logit and nested logit models

The BLP model

$$U_{ik} = X_k\beta_i - \alpha_i p_k + \xi_k + \varepsilon_{ik}$$
$$= \delta_k + \mu_{ik} + \varepsilon_{ik}$$

- X_k: observed product characteristics
- p_k: price of the product, endogenous: $\mathbb{E}(\xi|p) \neq 0$
- ξ_k: unobserved product characteristics (unobserved to the econometrician)
- δ_k: mean product utility
- μ_{ik}: individual deviation from the mean utility
- Important assumption: ξ_k are the same for all consumers, no unobserved heterogeneity in the valuation of the unobserved characteristics
- We examine first two restricted versions of the BLP model: logit and nested logit models

The BLP model

$$U_{ik} = X_k\beta_i - \alpha_i p_k + \xi_k + \varepsilon_{ik}$$
$$= \delta_k + \mu_{ik} + \varepsilon_{ik}$$

- X_k: observed product characteristics
- p_k: price of the product, endogenous: $\mathbb{E}(\xi|p) \neq 0$
- ξ_k: unobserved product characteristics (unobserved to the econometrician)
- δ_k: mean product utility
- μ_{ik}: individual deviation from the mean utility
- Important assumption: ξ_k are the same for all consumers, no unobserved heterogeneity in the valuation of the unobserved characteristics
- We examine first two restricted versions of the BLP model: logit and nested logit models

The BLP model

$$U_{ik} = X_k\beta_i - \alpha_i p_k + \xi_k + \varepsilon_{ik}$$
$$= \delta_k + \mu_{ik} + \varepsilon_{ik}$$

- X_k: observed product characteristics
- p_k: price of the product, endogenous: $\mathbb{E}(\xi|p) \neq 0$
- ξ_k: unobserved product characteristics (unobserved to the econometrician)
- δ_k: mean product utility
- μ_{ik}: individual deviation from the mean utility
- Important assumption: ξ_k are the same for all consumers, no unobserved heterogeneity in the valuation of the unobserved characteristics
- We examine first two restricted versions of the BLP model: logit and nested logit models

The BLP model

$$U_{ik} = X_k\beta_i - \alpha_i p_k + \xi_k + \varepsilon_{ik}$$
$$= \delta_k + \mu_{ik} + \varepsilon_{ik}$$

- X_k: observed product characteristics
- p_k: price of the product, endogenous: $\mathbb{E}(\xi|p) \neq 0$
- ξ_k: unobserved product characteristics (unobserved to the econometrician)
- δ_k: mean product utility
- μ_{ik}: individual deviation from the mean utility
- Important assumption: ξ_k are the same for all consumers, no unobserved heterogeneity in the valuation of the unobserved characteristics
- We examine first two restricted versions of the BLP model: logit and nested logit models

The BLP model

$$U_{ik} = X_k\beta_i - \alpha_i p_k + \xi_k + \varepsilon_{ik}$$
$$= \delta_k + \mu_{ik} + \varepsilon_{ik}$$

- X_k: observed product characteristics
- p_k: price of the product, endogenous: $\mathbb{E}(\xi|p) \neq 0$
- ξ_k: unobserved product characteristics (unobserved to the econometrician)
- δ_k: mean product utility
- μ_{ik}: individual deviation from the mean utility
- Important assumption: ξ_k are the same for all consumers, no unobserved heterogeneity in the valuation of the unobserved characteristics
- We examine first two restricted versions of the BLP model: logit and nested logit models

The Logit model

- Assume homogeneous preferences: $\alpha_i = \alpha$, $\beta_i = \beta\ \forall i$

$$s_k(\alpha, \beta) = \frac{\exp\left(X_k\beta - \alpha p_k + \xi_k\right)}{\sum_{j=0}^{J} \exp\left(X_j\beta - \alpha p_j + \xi_j\right)}$$

- Estimation of the model based on moment conditions (GMM):

$$\mathbb{E}(\xi | Z) = 0$$

- So we need a market share inversion to obtain $\xi(\alpha, \beta)$
- Trick: express the share of the "outside good", which mean utility is normalized to 0, $\delta_0 = 0$

$$s_0(\alpha, \beta) = \frac{1}{\sum_{j=0}^{J} \exp\left(X_j\beta - \alpha p_j + \xi_j\right)}$$

The Logit model

- Assume homogeneous preferences: $\alpha_i = \alpha$, $\beta_i = \beta \; \forall i$

$$s_k(\alpha, \beta) = \frac{\exp\left(X_k \beta - \alpha p_k + \xi_k\right)}{\sum_{j=0}^{J} \exp\left(X_j \beta - \alpha p_j + \xi_j\right)}$$

- Estimation of the model based on moment conditions (GMM):

$$\mathbb{E}(\xi | Z) = 0$$

- So we need a market share inversion to obtain $\xi(\alpha, \beta)$
- Trick: express the share of the "outside good", which mean utility is normalized to 0, $\delta_0 = 0$

$$s_0(\alpha, \beta) = \frac{1}{\sum_{j=0}^{J} \exp\left(X_j \beta - \alpha p_j + \xi_j\right)}$$

The Logit model

- Assume homogeneous preferences: $\alpha_i = \alpha$, $\beta_i = \beta \; \forall i$

$$s_k(\alpha, \beta) = \frac{\exp\left(X_k\beta - \alpha p_k + \xi_k\right)}{\sum_{j=0}^{J} \exp\left(X_j\beta - \alpha p_j + \xi_j\right)}$$

- Estimation of the model based on moment conditions (GMM):

$$\mathbb{E}(\xi | Z) = 0$$

- So we need a market share inversion to obtain $\xi(\alpha, \beta)$
- Trick: express the share of the "outside good", which mean utility is normalized to 0, $\delta_0 = 0$

$$s_0(\alpha, \beta) = \frac{1}{\sum_{j=0}^{J} \exp\left(X_j\beta - \alpha p_j + \xi_j\right)}$$

The Logit model

- Assume homogeneous preferences: $\alpha_i = \alpha$, $\beta_i = \beta \; \forall i$

$$s_k(\alpha, \beta) = \frac{\exp\left(X_k\beta - \alpha p_k + \xi_k\right)}{\sum_{j=0}^{J} \exp\left(X_j\beta - \alpha p_j + \xi_j\right)}$$

- Estimation of the model based on moment conditions (GMM):

$$\mathbb{E}(\xi | Z) = 0$$

- So we need a market share inversion to obtain $\xi(\alpha, \beta)$
- Trick: express the share of the "outside good", which mean utility is normalized to 0, $\delta_0 = 0$

$$s_0(\alpha, \beta) = \frac{1}{\sum_{j=0}^{J} \exp\left(X_j\beta - \alpha p_j + \xi_j\right)}$$

The Logit model

- Assume homogeneous preferences: $\alpha_i = \alpha$, $\beta_i = \beta \; \forall i$

$$s_k(\alpha, \beta) = \frac{\exp\left(X_k\beta - \alpha p_k + \xi_k\right)}{\sum_{j=0}^{J} \exp\left(X_j\beta - \alpha p_j + \xi_j\right)}$$

- Estimation of the model based on moment conditions (GMM):

$$\mathbb{E}(\xi | Z) = 0$$

- So we need a market share inversion to obtain $\xi(\alpha, \beta)$
- Trick: express the share of the "outside good", which mean utility is normalized to 0, $\delta_0 = 0$

$$s_0(\alpha, \beta) = \frac{1}{\sum_{j=0}^{J} \exp\left(X_j\beta - \alpha p_j + \xi_j\right)}$$

The Logit model

- Assume homogeneous preferences: $\alpha_i = \alpha$, $\beta_i = \beta\ \forall i$

$$s_k(\alpha,\beta) = \frac{\exp\left(X_k\beta - \alpha p_k + \xi_k\right)}{\sum_{j=0}^{J}\exp\left(X_j\beta - \alpha p_j + \xi_j\right)}$$

- Estimation of the model based on moment conditions (GMM):

$$\mathbb{E}(\xi|Z) = 0$$

- So we need a market share inversion to obtain $\xi(\alpha,\beta)$

- Trick: express the share of the "outside good", which mean utility is normalized to 0, $\delta_0 = 0$

$$s_0(\alpha,\beta) = \frac{1}{\sum_{j=0}^{J}\exp\left(X_j\beta - \alpha p_j + \xi_j\right)}$$

The Logit model

- Assume homogeneous preferences: $\alpha_i = \alpha$, $\beta_i = \beta \; \forall i$

$$s_k(\alpha, \beta) = \frac{\exp\left(X_k\beta - \alpha p_k + \xi_k\right)}{\sum_{j=0}^{J} \exp\left(X_j\beta - \alpha p_j + \xi_j\right)}$$

- Estimation of the model based on moment conditions (GMM):

$$\mathbb{E}(\xi | Z) = 0$$

- So we need a market share inversion to obtain $\xi(\alpha, \beta)$

- Trick: express the share of the "outside good", which mean utility is normalized to 0, $\delta_0 = 0$

$$s_0(\alpha, \beta) = \frac{1}{\sum_{j=0}^{J} \exp\left(X_j\beta - \alpha p_j + \xi_j\right)}$$

The Logit model

- Assume homogeneous preferences: $\alpha_i = \alpha$, $\beta_i = \beta \; \forall i$

$$s_k(\alpha, \beta) = \frac{\exp\left(X_k \beta - \alpha p_k + \xi_k\right)}{\sum_{j=0}^{J} \exp\left(X_j \beta - \alpha p_j + \xi_j\right)}$$

- Estimation of the model based on moment conditions (GMM):

$$\mathbb{E}(\xi | Z) = 0$$

- So we need a market share inversion to obtain $\xi(\alpha, \beta)$
- Trick: express the share of the “outside good”, which mean utility is normalized to 0, $\delta_0 = 0$

$$s_0(\alpha, \beta) = \frac{1}{\sum_{j=0}^{J} \exp\left(X_j \beta - \alpha p_j + \xi_j\right)}$$

The Logit model

- Assume homogeneous preferences: $\alpha_i = \alpha$, $\beta_i = \beta \; \forall i$

$$s_k(\alpha, \beta) = \frac{\exp\left(X_k \beta - \alpha p_k + \xi_k\right)}{\sum_{j=0}^{J} \exp\left(X_j \beta - \alpha p_j + \xi_j\right)}$$

- Estimation of the model based on moment conditions (GMM):

$$\mathbb{E}(\xi | Z) = 0$$

- So we need a market share inversion to obtain $\xi(\alpha, \beta)$
- Trick: express the share of the "outside good", which mean utility is normalized to 0, $\delta_0 = 0$

$$s_0(\alpha, \beta) = \frac{1}{\sum_{j=0}^{J} \exp\left(X_j \beta - \alpha p_j + \xi_j\right)}$$

The Logit model

- Assume homogeneous preferences: $\alpha_i = \alpha$, $\beta_i = \beta \; \forall i$
$$s_k(\alpha, \beta) = \frac{\exp\left(X_k\beta - \alpha p_k + \xi_k\right)}{\sum_{j=0}^{J} \exp\left(X_j\beta - \alpha p_j + \xi_j\right)}$$
- Estimation of the model based on moment conditions (GMM):
$$\mathbb{E}(\xi|Z) = 0$$
- So we need a market share inversion to obtain $\xi(\alpha, \beta)$
- Trick: express the share of the "outside good", which mean utility is normalized to 0, $\delta_0 = 0$
$$s_0(\alpha, \beta) = \frac{1}{\sum_{j=0}^{J} \exp\left(X_j\beta - \alpha p_j + \xi_j\right)}$$

The Logit model

So we can invert market shares and obtain $\xi(\alpha, \beta)$ as:

$$\xi_k(\alpha, \beta) = \ln(s_k) - \ln(s_0) - X_k\beta - \alpha p_k$$

We have a linear relation so we can use 2SLS, need instruments for the price
Instruments:

- Cost shifters
- Functions of other product characteristics (e.g. $\sum_{j \neq k} X_j$), intuition = a product's margin and price is constrained by the existence of close substitutes. Function of other products' characteristics as proxy for how crowded is the product space
- Prices in other independent markets (e.g. different cities), prices in different cities correlated through the costs but exogeneity may be problematic if national advertising campaign

The Logit model

So we can invert market shares and obtain $\xi(\alpha, \beta)$ as:

$$\xi_k(\alpha, \beta) = \ln(s_k) - \ln(s_0) - X_k\beta - \alpha p_k$$

We have a linear relation so we can use 2SLS, need instruments for the price

Instruments:

- Cost shifters
- Functions of other product characteristics (e.g. $\sum_{j \neq k} X_j$), intuition = a product's margin and price is constrained by the existence of close substitutes. Function of other products' characteristics as proxy for how crowded is the product space
- Prices in other independent markets (e.g. different cities), prices in different cities correlated through the costs but exogeneity may be problematic if national advertising campaign

The Logit model

So we can invert market shares and obtain $\xi(\alpha, \beta)$ as:

$$\xi_k(\alpha, \beta) = \ln(s_k) - \ln(s_0) - X_k\beta - \alpha p_k$$

We have a linear relation so we can use 2SLS, need instruments for the price
Instruments:

- Cost shifters
- Functions of other product characteristics (e.g. $\sum_{j \neq k} X_j$), intuition = a product's margin and price is constrained by the existence of close substitutes. Function of other products' characteristics as proxy for how crowded is the product space
- Prices in other independent markets (e.g. different cities), prices in different cities correlated through the costs but exogeneity may be problematic if national advertising campaign

The Logit model

So we can invert market shares and obtain $\xi(\alpha, \beta)$ as:

$$\xi_k(\alpha, \beta) = \ln(s_k) - \ln(s_0) - X_k\beta - \alpha p_k$$

We have a linear relation so we can use 2SLS, need instruments for the price
Instruments:

- Cost shifters
- Functions of other product characteristics (e.g. $\sum_{j \neq k} X_j$), intuition = a product's margin and price is constrained by the existence of close substitutes. Function of other products' characteristics as proxy for how crowded is the product space
- Prices in other independent markets (e.g. different cities), prices in different cities correlated through the costs but exogeneity may be problematic if national advertising campaign

The Logit model

So we can invert market shares and obtain $\xi(\alpha, \beta)$ as:

$$\xi_k(\alpha, \beta) = \ln(s_k) - \ln(s_0) - X_k\beta - \alpha p_k$$

We have a linear relation so we can use 2SLS, need instruments for the price

Instruments:

- Cost shifters
- Functions of other product characteristics (e.g. $\sum_{j \neq k} X_j$), intuition = a product's margin and price is constrained by the existence of close substitutes. Function of other products' characteristics as proxy for how crowded is the product space
- Prices in other independent markets (e.g. different cities), prices in different cities correlated through the costs but exogeneity may be problematic if national advertising campaign

Firm Behavior

- Profits:

$$\pi_f = \sum_{k \in J_f} [p_k - mc_k] \cdot M \cdot s_k(p, x, \xi; \alpha, \beta)$$

where M denotes the number of consumers in the market

- As before, use FOC for optimal prices:

$$M \cdot s_k(p, x, \xi; \alpha, \beta) + \sum_{l \in J_f} [p_l - mc_l] \cdot M \cdot \frac{\partial s_l(p, x, \xi; \alpha, \beta)}{\partial p_k} = 0$$

- For a single product firm:
  $mc_k = p_k + s_k(p, x, \xi; \alpha, \beta) / \frac{\partial s_k(p, x, \xi; \alpha, \beta)}{\partial p_k}$
- Specification of marginal cost equation:
  $mc_k = w_k\gamma + u_k$where w_k are cost shifters, and u_k is an unobserved cost error.

Firm Behavior

- Profits:

$$\pi_f = \sum_{k \in J_f} [p_k - mc_k] \cdot M \cdot s_k\left(p, x, \xi; \alpha, \beta\right)$$

where M denotes the number of consumers in the market

- As before, use FOC for optimal prices:

$$M \cdot s_k\left(p, x, \xi; \alpha, \beta\right) + \sum_{l \in J_f} [p_l - mc_l] \cdot M \cdot \frac{\partial s_l\left(p, x, \xi; \alpha, \beta\right)}{\partial p_k} = 0$$

- For a single product firm:
$mc_k = p_k + s_k\left(p, x, \xi; \alpha, \beta\right) / \frac{\partial s_k(p,x,\xi;\alpha,\beta)}{\partial p_k}$
- Specification of marginal cost equation:
$mc_k = w_k\gamma + u_k$where w_k are cost shifters, and u_k is an unobserved cost error.

Firm Behavior

- Profits:

$$\pi_f = \sum_{k \in J_f} [p_k - mc_k] \cdot M \cdot s_k(p, x, \xi; \alpha, \beta)$$

where M denotes the number of consumers in the market

- As before, use FOC for optimal prices:

$$M \cdot s_k(p, x, \xi; \alpha, \beta) + \sum_{l \in J_f} [p_l - mc_l] \cdot M \cdot \frac{\partial s_l(p, x, \xi; \alpha, \beta)}{\partial p_k} = 0$$

- For a single product firm:
  $mc_k = p_k + s_k(p, x, \xi; \alpha, \beta) / \frac{\partial s_k(p,x,\xi;\alpha,\beta)}{\partial p_k}$
- Specification of marginal cost equation:
  $mc_k = w_k\gamma + u_k$ where w_k are cost shifters, and u_k is an unobserved cost error.

Firm Behavior

- Profits:

$$\pi_f = \sum_{k \in J_f} [p_k - mc_k] \cdot M \cdot s_k(p, x, \xi; \alpha, \beta)$$

  where M denotes the number of consumers in the market
- As before, use FOC for optimal prices:

$$M \cdot s_k(p, x, \xi; \alpha, \beta) + \sum_{l \in J_f} [p_l - mc_l] \cdot M \cdot \frac{\partial s_l(p, x, \xi; \alpha, \beta)}{\partial p_k} = 0$$

- For a single product firm:
  $mc_k = p_k + s_k(p, x, \xi; \alpha, \beta) / \frac{\partial s_k(p,x,\xi;\alpha,\beta)}{\partial p_k}$
- Specification of marginal cost equation:
  $mc_k = w_k\gamma + u_k$where w_k are cost shifters, and u_k is an unobserved cost error.

Firm Behavior

- Profits:

$$\pi_f = \sum_{k \in J_f} [p_k - mc_k] \cdot M \cdot s_k (p, x, \xi; \alpha, \beta)$$

  where M denotes the number of consumers in the market
- As before, use FOC for optimal prices:

$$M \cdot s_k (p, x, \xi; \alpha, \beta) + \sum_{l \in J_f} [p_l - mc_l] \cdot M \cdot \frac{\partial s_l (p, x, \xi; \alpha, \beta)}{\partial p_k} = 0$$

- For a single product firm:
  $mc_k = p_k + s_k (p, x, \xi; \alpha, \beta) / \frac{\partial s_k (p, x, \xi; \alpha, \beta)}{\partial p_k}$
- Specification of marginal cost equation:
  $mc_k = w_k \gamma + u_k$ where w_k are cost shifters, and u_k is an unobserved cost error.

Firm Behavior

- Profits:

$$\pi_f = \sum_{k \in J_f} [p_k - mc_k] \cdot M \cdot s_k(p, x, \xi; \alpha, \beta)$$

where M denotes the number of consumers in the market

- As before, use FOC for optimal prices:

$$M \cdot s_k(p, x, \xi; \alpha, \beta) + \sum_{l \in J_f} [p_l - mc_l] \cdot M \cdot \frac{\partial s_l(p, x, \xi; \alpha, \beta)}{\partial p_k} = 0$$

- For a single product firm:
  $mc_k = p_k + s_k(p, x, \xi; \alpha, \beta) / \frac{\partial s_k(p,x,\xi;\alpha,\beta)}{\partial p_k}$

- Specification of marginal cost equation:
  $mc_k = w_k \gamma + u_k$ where w_k are cost shifters, and u_k is an unobserved cost error.

Firm Behavior

- Profits:

$$\pi_f = \sum_{k \in J_f} [p_k - mc_k] \cdot M \cdot s_k(p, x, \xi; \alpha, \beta)$$

  where M denotes the number of consumers in the market
- As before, use FOC for optimal prices:

$$M \cdot s_k(p, x, \xi; \alpha, \beta) + \sum_{l \in J_f} [p_l - mc_l] \cdot M \cdot \frac{\partial s_l(p, x, \xi; \alpha, \beta)}{\partial p_k} = 0$$

- For a single product firm:
  $mc_k = p_k + s_k(p, x, \xi; \alpha, \beta) / \frac{\partial s_k(p,x,\xi;\alpha,\beta)}{\partial p_k}$
- Specification of marginal cost equation:
  $mc_k = w_k\gamma + u_k$where w_k are cost shifters, and u_k is an unobserved cost error.

Limitations

- Price elasticities:

$$\frac{\partial s_k}{\partial p_k} = -\alpha \frac{\partial s_k}{\partial \delta_k} = -\alpha s_k(1 - s_k), \qquad \frac{\partial s_k}{\partial p_j} = \alpha s_k s_j$$

- Substitution between two products only depend on market shares
- We would prefer substitution to depend on characteristics of products: the more similar, the more substitute
- IIA (independence from irrelevant alternatives): the ratio of market shares of two products does not depend on other products $\frac{s_j}{s_k} = \frac{\exp(X_j\beta - \alpha p_j + \xi_j)}{\exp(X_k\beta - \alpha p_k + \xi_k)}$
- Solutions:
  - Nested Logit model
  - Random coefficients model

Limitations

- Price elasticities:

$$\frac{\partial s_k}{\partial p_k} = -\alpha \frac{\partial s_k}{\partial \delta_k} = -\alpha s_k(1 - s_k), \qquad \frac{\partial s_k}{\partial p_j} = \alpha s_k s_j$$

- Substitution between two products only depend on market shares
- We would prefer substitution to depend on characteristics of products: the more similar, the more substitute
- IIA (independence from irrelevant alternatives): the ratio of market shares of two products does not depend on other products $\frac{s_j}{s_k} = \frac{\exp(X_j\beta - \alpha p_j + \xi_j)}{\exp(X_k\beta - \alpha p_k + \xi_k)}$
- Solutions:
  - Nested Logit model
  - Random coefficients model

Limitations

- Price elasticities:

$$\frac{\partial s_k}{\partial p_k} = -\alpha \frac{\partial s_k}{\partial \delta_k} = -\alpha s_k(1 - s_k), \qquad \frac{\partial s_k}{\partial p_j} = \alpha s_k s_j$$

- Substitution between two products only depend on market shares
- We would prefer substitution to depend on characteristics of products: the more similar, the more substitute
- IIA (independence from irrelevant alternatives): the ratio of market shares of two products does not depend on other products $\frac{s_j}{s_k} = \frac{\exp(X_j\beta - \alpha p_j + \xi_j)}{\exp(X_k\beta - \alpha p_k + \xi_k)}$
- Solutions:
    - Nested Logit model
    - Random coefficients model

Limitations

- Price elasticities:

$$\frac{\partial s_k}{\partial p_k} = -\alpha \frac{\partial s_k}{\partial \delta_k} = -\alpha s_k(1 - s_k), \qquad \frac{\partial s_k}{\partial p_j} = \alpha s_k s_j$$

- Substitution between two products only depend on market shares
- We would prefer substitution to depend on characteristics of products: the more similar, the more substitute
- IIA (independence from irrelevant alternatives): the ratio of market shares of two products does not depend on other products $\frac{s_j}{s_k} = \frac{\exp(X_j\beta - \alpha p_j + \xi_j)}{\exp(X_k\beta - \alpha p_k + \xi_k)}$
- Solutions:
  - Nested Logit model
  - Random coefficients model

Limitations

- Price elasticities:

$$\frac{\partial s_k}{\partial p_k} = -\alpha \frac{\partial s_k}{\partial \delta_k} = -\alpha s_k(1 - s_k), \qquad \frac{\partial s_k}{\partial p_j} = \alpha s_k s_j$$

- Substitution between two products only depend on market shares
- We would prefer substitution to depend on characteristics of products: the more similar, the more substitute
- IIA (independence from irrelevant alternatives): the ratio of market shares of two products does not depend on other products $\frac{s_j}{s_k} = \frac{\exp(X_j\beta - \alpha p_j + \xi_j)}{\exp(X_k\beta - \alpha p_k + \xi_k)}$
- Solutions:
  - Nested Logit model
  - Random coefficients model

Limitations

- Price elasticities:

$$\frac{\partial s_k}{\partial p_k} = -\alpha \frac{\partial s_k}{\partial \delta_k} = -\alpha s_k(1 - s_k), \qquad \frac{\partial s_k}{\partial p_j} = \alpha s_k s_j$$

- Substitution between two products only depend on market shares
- We would prefer substitution to depend on characteristics of products: the more similar, the more substitute
- IIA (independence from irrelevant alternatives): the ratio of market shares of two products does not depend on other products $\frac{s_j}{s_k} = \frac{\exp(X_j\beta - \alpha p_j + \xi_j)}{\exp(X_k\beta - \alpha p_k + \xi_k)}$
- Solutions:
  - Nested Logit model
  - Random coefficients model

Limitations

- Price elasticities:

$$\frac{\partial s_k}{\partial p_k} = -\alpha \frac{\partial s_k}{\partial \delta_k} = -\alpha s_k(1 - s_k), \qquad \frac{\partial s_k}{\partial p_j} = \alpha s_k s_j$$

- Substitution between two products only depend on market shares
- We would prefer substitution to depend on characteristics of products: the more similar, the more substitute
- IIA (independence from irrelevant alternatives): the ratio of market shares of two products does not depend on other products $\frac{s_j}{s_k} = \frac{\exp\left(X_j\beta - \alpha p_j + \xi_j\right)}{\exp(X_k\beta - \alpha p_k + \xi_k)}$
- Solutions:
  - Nested Logit model
  - Random coefficients model

Limitations

- Price elasticities:

$$\frac{\partial s_k}{\partial p_k} = -\alpha \frac{\partial s_k}{\partial \delta_k} = -\alpha s_k(1 - s_k), \qquad \frac{\partial s_k}{\partial p_j} = \alpha s_k s_j$$

- Substitution between two products only depend on market shares
- We would prefer substitution to depend on characteristics of products: the more similar, the more substitute
- IIA (independence from irrelevant alternatives): the ratio of market shares of two products does not depend on other products $\frac{s_j}{s_k} = \frac{\exp\left(X_j\beta - \alpha p_j + \xi_j\right)}{\exp(X_k\beta - \alpha p_k + \xi_k)}$

Limitations

- Price elasticities:

$$\frac{\partial s_k}{\partial p_k} = -\alpha \frac{\partial s_k}{\partial \delta_k} = -\alpha s_k(1 - s_k), \qquad \frac{\partial s_k}{\partial p_j} = \alpha s_k s_j$$

- Substitution between two products only depend on market shares
- We would prefer substitution to depend on characteristics of products: the more similar, the more substitute
- IIA (independence from irrelevant alternatives): the ratio of market shares of two products does not depend on other products $\frac{s_j}{s_k} = \frac{\exp\left(X_j\beta - \alpha p_j + \xi_j\right)}{\exp(X_k\beta - \alpha p_k + \xi_k)}$
- Solutions:
  - Nested Logit model
  - Random coefficients model

Limitations

- Price elasticities:

$$\frac{\partial s_k}{\partial p_k} = -\alpha \frac{\partial s_k}{\partial \delta_k} = -\alpha s_k(1 - s_k), \qquad \frac{\partial s_k}{\partial p_j} = \alpha s_k s_j$$

- Substitution between two products only depend on market shares
- We would prefer substitution to depend on characteristics of products: the more similar, the more substitute
- IIA (independence from irrelevant alternatives): the ratio of market shares of two products does not depend on other products $\frac{s_j}{s_k} = \frac{\exp\left(X_j\beta - \alpha p_j + \xi_j\right)}{\exp(X_k\beta - \alpha p_k + \xi_k)}$
- Solutions:
  - Nested Logit model
  - Random coefficients model

Limitations

- Price elasticities:

$$\frac{\partial s_k}{\partial p_k} = -\alpha \frac{\partial s_k}{\partial \delta_k} = -\alpha s_k(1 - s_k), \qquad \frac{\partial s_k}{\partial p_j} = \alpha s_k s_j$$

- Substitution between two products only depend on market shares
- We would prefer substitution to depend on characteristics of products: the more similar, the more substitute
- IIA (independence from irrelevant alternatives): the ratio of market shares of two products does not depend on other products $\frac{s_j}{s_k} = \frac{\exp\left(X_j\beta - \alpha p_j + \xi_j\right)}{\exp(X_k\beta - \alpha p_k + \xi_k)}$
- Solutions:
  - Nested Logit model
  - Random coefficients model

Limitations

- Price elasticities:

$$\frac{\partial s_k}{\partial p_k} = -\alpha \frac{\partial s_k}{\partial \delta_k} = -\alpha s_k(1 - s_k), \qquad \frac{\partial s_k}{\partial p_j} = \alpha s_k s_j$$

- Substitution between two products only depend on market shares
- We would prefer substitution to depend on characteristics of products: the more similar, the more substitute
- IIA (independence from irrelevant alternatives): the ratio of market shares of two products does not depend on other products $\frac{s_j}{s_k} = \frac{\exp\left(X_j\beta - \alpha p_j + \xi_j\right)}{\exp(X_k\beta - \alpha p_k + \xi_k)}$
- Solutions:
  - Nested Logit model
  - Random coefficients model

Nested Logit Model

Idea: Group products that are more similar and assume common group-specific individual taste shock

Model consistent with a sequential decision: (i) the group (or segment) and (ii) the product inside the group

Example of segmentation:

- City car, sports car, family car...
- Choice between credit card or debit card and then which bank
- Or first the choice of a bank and then the choice of a payment card

Notes:

- Definition of segments not unique!
- We can try different segmentation to find the most relevant one
- We can have multiple layers of nests
- IIA still holds for products in the same nest but not for products in different nests

Nested Logit Model

Idea: Group products that are more similar and assume common group-specific individual taste shock

Model consistent with a sequential decision: (i) the group (or segment) and (ii) the product inside the group

Example of segmentation:

- City car, sports car, family car...
- Choice between credit card or debit card and then which bank
- Or first the choice of a bank and then the choice of a payment card

Notes:

- Definition of segments not unique!
- We can try different segmentation to find the most relevant one
- We can have multiple layers of nests
- IIA still holds for products in the same nest but not for products in different nests

Nested Logit Model

Idea: Group products that are more similar and assume common group-specific individual taste shock

Model consistent with a sequential decision: (i) the group (or segment) and (ii) the product inside the group

Example of segmentation:

- City car, sports car, family car...
- Choice between credit card or debit card and then which bank
- Or first the choice of a bank and then the choice of a payment card

Notes:

- Definition of segments not unique!
- We can try different segmentation to find the most relevant one
- We can have multiple layers of nests
- IIA still holds for products in the same nest but not for products in different nests

Nested Logit Model

Idea: Group products that are more similar and assume common group-specific individual taste shock

Model consistent with a sequential decision: (i) the group (or segment) and (ii) the product inside the group

Example of segmentation:

- City car, sports car, family car...
- Choice between credit card or debit card and then which bank
- Or first the choice of a bank and then the choice of a payment card

Notes:

- Definition of segments not unique!
- We can try different segmentation to find the most relevant one
- We can have multiple layers of nests
- IIA still holds for products in the same nest but not for products in different nests

Nested Logit Model

Idea: Group products that are more similar and assume common group-specific individual taste shock

Model consistent with a sequential decision: (i) the group (or segment) and (ii) the product inside the group

Example of segmentation:

- City car, sports car, family car...
- Choice between credit card or debit card and then which bank
- Or first the choice of a bank and then the choice of a payment card

Notes:

- Definition of segments not unique!
- We can try different segmentation to find the most relevant one
- We can have multiple layers of nests
- IIA still holds for products in the same nest but not for products in different nests

Nested Logit Model

Idea: Group products that are more similar and assume common group-specific individual taste shock

Model consistent with a sequential decision: (i) the group (or segment) and (ii) the product inside the group

Example of segmentation:

- City car, sports car, family car...
- Choice between credit card or debit card and then which bank
- Or first the choice of a bank and then the choice of a payment card

Notes:

- Definition of segments not unique!
- We can try different segmentation to find the most relevant one
- We can have multiple layers of nests
- IIA still holds for products in the same nest but not for products in different nests

Nested Logit Model

Idea: Group products that are more similar and assume common group-specific individual taste shock

Model consistent with a sequential decision: (i) the group (or segment) and (ii) the product inside the group

Example of segmentation:

- City car, sports car, family car...
- Choice between credit card or debit card and then which bank
- Or first the choice of a bank and then the choice of a payment card

Notes:

- Definition of segments not unique!
- We can try different segmentation to find the most relevant one
- We can have multiple layers of nests
- IIA still holds for products in the same nest but not for products in different nests

Nested Logit Model

Idea: Group products that are more similar and assume common group-specific individual taste shock

Model consistent with a sequential decision: (i) the group (or segment) and (ii) the product inside the group

Example of segmentation:

- City car, sports car, family car...
- Choice between credit card or debit card and then which bank
- Or first the choice of a bank and then the choice of a payment card

Notes:

- Definition of segments not unique!
- We can try different segmentation to find the most relevant one
- We can have multiple layers of nests
- IIA still holds for products in the same nest but not for products in different nests

Nested Logit Model

Idea: Group products that are more similar and assume common group-specific individual taste shock

Model consistent with a sequential decision: (i) the group (or segment) and (ii) the product inside the group

Example of segmentation:

- City car, sports car, family car...
- Choice between credit card or debit card and then which bank
- Or first the choice of a bank and then the choice of a payment card

Notes:

- Definition of segments not unique!
- We can try different segmentation to find the most relevant one
- We can have multiple layers of nests
- IIA still holds for products in the same nest but not for products in different nests

Nested Logit Model

Idea: Group products that are more similar and assume common group-specific individual taste shock

Model consistent with a sequential decision: (i) the group (or segment) and (ii) the product inside the group

Example of segmentation:

- City car, sports car, family car...
- Choice between credit card or debit card and then which bank
- Or first the choice of a bank and then the choice of a payment card

Notes:

- Definition of segments not unique!
- We can try different segmentation to find the most relevant one
- We can have multiple layers of nests
- IIA still holds for products in the same nest but not for products in different nests

Nested Logit Model

Idea: Group products that are more similar and assume common group-specific individual taste shock

Model consistent with a sequential decision: (i) the group (or segment) and (ii) the product inside the group

Example of segmentation:

- City car, sports car, family car...
- Choice between credit card or debit card and then which bank
- Or first the choice of a bank and then the choice of a payment card

Notes:

- Definition of segments not unique!
- We can try different segmentation to find the most relevant one
- We can have multiple layers of nests
- IIA still holds for products in the same nest but not for products in different nests

Nested Logit Model

Formalization:

$$\varepsilon_{ij} = \zeta_{ig} + (1 - \sigma)\epsilon_{ij}$$

- ϵ_{ij} is iid extreme value, ζ_{ig} distributed such that ε_{ij} is also extreme value
- σ represents the intra-group degree of substitution, it belongs to $[0, 1]$
- If $\sigma \to 0$: logit case, groups are irrelevant
- If $\sigma \to 1$: Substitution inside group only

Remark:

- We can use σ to test relevance of the segmentation
  H_0: $\sigma = 0$ vs. H_a: $\sigma \neq 0$
- With multiple levels of nests, we should expect the lower nests to have higher correlations, we can also use this restriction to test the relevance of the ordering of the nests:
  H_0: $\sigma_l \geq \sigma_h$ vs. H_a: $\sigma_l < \sigma_h$

Nested Logit Model

Formalization:

$$\varepsilon_{ij} = \zeta_{ig} + (1 - \sigma)\epsilon_{ij}$$

- ϵ_{ij} is iid extreme value, ζ_{ig} distributed such that ε_{ij} is also extreme value
- σ represents the intra-group degree of substitution, it belongs to $[0, 1]$
- If $\sigma \to 0$: logit case, groups are irrelevant
- If $\sigma \to 1$: Substitution inside group only

Remark:

- We can use σ to test relevance of the segmentation H_0: $\sigma = 0$ vs. H_a: $\sigma \neq 0$
- With multiple levels of nests, we should expect the lower nests to have higher correlations, we can also use this restriction to test the relevance of the ordering of the nests: H_0: $\sigma_l \geq \sigma_h$ vs. H_a: $\sigma_l < \sigma_h$

Nested Logit Model

Formalization:

$$\varepsilon_{ij} = \zeta_{ig} + (1 - \sigma)\epsilon_{ij}$$

- ϵ_{ij} is iid extreme value, ζ_{ig} distributed such that ε_{ij} is also extreme value
- σ represents the intra-group degree of substitution, it belongs to $[0, 1]$
- If $\sigma \to 0$: logit case, groups are irrelevant
- If $\sigma \to 1$: Substitution inside group only

Remark:

- We can use σ to test relevance of the segmentation H_0: $\sigma = 0$ vs. H_a: $\sigma \neq 0$
- With multiple levels of nests, we should expect the lower nests to have higher correlations, we can also use this restriction to test the relevance of the ordering of the nests: H_0: $\sigma_l \geq \sigma_h$ vs. H_a: $\sigma_l < \sigma_h$

Nested Logit Model

Formalization:

$$\varepsilon_{ij} = \zeta_{ig} + (1 - \sigma)\epsilon_{ij}$$

- ϵ_{ij} is iid extreme value, ζ_{ig} distributed such that ε_{ij} is also extreme value
- σ represents the intra-group degree of substitution, it belongs to $[0, 1]$
- If $\sigma \to 0$: logit case, groups are irrelevant
- If $\sigma \to 1$: Substitution inside group only

Remark:

- We can use σ to test relevance of the segmentation H_0: $\sigma = 0$ vs. H_a: $\sigma \neq 0$
- With multiple levels of nests, we should expect the lower nests to have higher correlations, we can also use this restriction to test the relevance of the ordering of the nests: H_0: $\sigma_l \geq \sigma_h$ vs. H_a: $\sigma_l < \sigma_h$

Nested Logit Model

Formalization:

$$\varepsilon_{ij} = \zeta_{ig} + (1-\sigma)\epsilon_{ij}$$

- ϵ_{ij} is iid extreme value, ζ_{ig} distributed such that ε_{ij} is also extreme value
- σ represents the intra-group degree of substitution, it belongs to $[0, 1]$
- If $\sigma \to 0$: logit case, groups are irrelevant
- If $\sigma \to 1$: Substitution inside group only

Remark:

- We can use σ to test relevance of the segmentation H_0: $\sigma = 0$ vs. H_a: $\sigma \neq 0$
- With multiple levels of nests, we should expect the lower nests to have higher correlations, we can also use this restriction to test the relevance of the ordering of the nests: H_0: $\sigma_l \geq \sigma_h$ vs. H_a: $\sigma_l < \sigma_h$

Nested Logit Model

Formalization:

$$\varepsilon_{ij} = \zeta_{ig} + (1 - \sigma)\epsilon_{ij}$$

- ϵ_{ij} is iid extreme value, ζ_{ig} distributed such that ε_{ij} is also extreme value
- σ represents the intra-group degree of substitution, it belongs to $[0, 1]$
- If $\sigma \to 0$: logit case, groups are irrelevant
- If $\sigma \to 1$: Substitution inside group only

Remark:

- We can use σ to test relevance of the segmentation
  H_0: $\sigma = 0$ vs. H_a: $\sigma \neq 0$
- With multiple levels of nests, we should expect the lower nests to have higher correlations, we can also use this restriction to test the relevance of the ordering of the nests:
  H_0: $\sigma_l \geq \sigma_h$ vs. H_a: $\sigma_l < \sigma_h$

Nested Logit Model

Formalization:

$$\varepsilon_{ij} = \zeta_{ig} + (1 - \sigma)\epsilon_{ij}$$

- ϵ_{ij} is iid extreme value, ζ_{ig} distributed such that ε_{ij} is also extreme value
- σ represents the intra-group degree of substitution, it belongs to $[0, 1]$
- If $\sigma \to 0$: logit case, groups are irrelevant
- If $\sigma \to 1$: Substitution inside group only

Remark:

- We can use σ to test relevance of the segmentation
  H_0: $\sigma = 0$ vs. H_a: $\sigma \neq 0$
- With multiple levels of nests, we should expect the lower nests to have higher correlations, we can also use this restriction to test the relevance of the ordering of the nests:
  H_0: $\sigma_l \geq \sigma_h$ vs. H_a: $\sigma_l < \sigma_h$

Nested Logit Model

$$U_{ij} = X_j\beta_i - \alpha_i p_j + \xi_j + \zeta_{ig} + (1-\sigma)\epsilon_{ij}$$

Define $\mathcal{J}_g$ as the set of products in nest g.

Within nest market shares: ζ_{ig} irrelevant for the within-nest choice

$$\begin{aligned} s_{j|g} &= P(U_{ij} \geq U_{ik} \forall k \in \mathcal{J}_g) \\ &= P(\delta_j + (1-\sigma)\epsilon_{ij} \geq \delta_k + (1-\sigma)\epsilon_{ik} \forall k \in \mathcal{J}_g) \\ &= P(\tfrac{\delta_j}{1-\sigma} + \epsilon_{ij} \geq \tfrac{\delta_k}{1-\sigma} + \epsilon_{ik} \forall k \in \mathcal{J}_g) \end{aligned}$$

Same as logit probability, except $\delta \rightarrow \delta/(1-\sigma)$

$$s_{j|g} = \frac{\exp\left(\delta_j/(1-\sigma)\right)}{\sum_{k\in\mathcal{J}_g} \exp\left(\delta_k/(1-\sigma)\right)}$$

Nested Logit Model

$$U_{ij} = X_j\beta_i - \alpha_i p_j + \xi_j + \zeta_{ig} + (1-\sigma)\epsilon_{ij}$$

Define $\mathcal{J}_g$ as the set of products in nest g.

Within nest market shares: ζ_{ig} irrelevant for the within-nest choice

$$\begin{aligned} s_{j|g} &= P(U_{ij} \geq U_{ik} \forall k \in \mathcal{J}_g) \\ &= P(\delta_j + (1-\sigma)\epsilon_{ij} \geq \delta_k + (1-\sigma)\epsilon_{ik} \forall k \in \mathcal{J}_g) \\ &= P(\tfrac{\delta_j}{1-\sigma} + \epsilon_{ij} \geq \tfrac{\delta_k}{1-\sigma} + \epsilon_{ik} \forall k \in \mathcal{J}_g) \end{aligned}$$

Same as logit probability, except $\delta \to \delta/(1-\sigma)$

$$s_{j|g} = \frac{\exp\left(\delta_j/(1-\sigma)\right)}{\sum_{k \in \mathcal{J}_g} \exp\left(\delta_k/(1-\sigma)\right)}$$

Nested Logit Model

$$U_{ij} = X_j\beta_i - \alpha_i p_j + \xi_j + \zeta_{ig} + (1-\sigma)\epsilon_{ij}$$

Define $\mathcal{J}_g$ as the set of products in nest g.

Within nest market shares: ζ_{ig} irrelevant for the within-nest choice

$$\begin{aligned} s_{j|g} &= P(U_{ij} \geq U_{ik} \forall k \in \mathcal{J}_g) \\ &= P(\delta_j + (1-\sigma)\epsilon_{ij} \geq \delta_k + (1-\sigma)\epsilon_{ik} \forall k \in \mathcal{J}_g) \\ &= P(\tfrac{\delta_j}{1-\sigma} + \epsilon_{ij} \geq \tfrac{\delta_k}{1-\sigma} + \epsilon_{ik} \forall k \in \mathcal{J}_g) \end{aligned}$$

Same as logit probability, except $\delta \to \delta/(1-\sigma)$

$$s_{j|g} = \frac{\exp\left(\delta_j/(1-\sigma)\right)}{\sum_{k\in\mathcal{J}_g} \exp\left(\delta_k/(1-\sigma)\right)}$$

Nested Logit Model

$$U_{ij} = X_j\beta_i - \alpha_i p_j + \xi_j + \zeta_{ig} + (1-\sigma)\epsilon_{ij}$$

Define $\mathcal{J}_g$ as the set of products in nest g.

Within nest market shares: ζ_{ig} irrelevant for the within-nest choice

$$\begin{aligned} s_{j|g} &= P(U_{ij} \geq U_{ik} \forall k \in \mathcal{J}_g) \\ &= P(\delta_j + (1-\sigma)\epsilon_{ij} \geq \delta_k + (1-\sigma)\epsilon_{ik} \forall k \in \mathcal{J}_g) \\ &= P(\tfrac{\delta_j}{1-\sigma} + \epsilon_{ij} \geq \tfrac{\delta_k}{1-\sigma} + \epsilon_{ik} \forall k \in \mathcal{J}_g) \end{aligned}$$

Same as logit probability, except $\delta \to \delta/(1-\sigma)$

$$s_{j|g} = \frac{\exp\left(\delta_j/(1-\sigma)\right)}{\sum_{k\in\mathcal{J}_g} \exp\left(\delta_k/(1-\sigma)\right)}$$

Nested Logit Model

We define I_g as the inclusive value of nest g, i.e. the expected utility of the best product of the nest:

$$I_g = (1-\sigma)\ln \sum_{k \in \mathcal{J}_g} \exp(\delta_k/(1-\sigma))$$

Choice of a nest s_g:

$$\begin{aligned}
\bar{s}_g &= \frac{\exp I_g}{\sum_{g'=0}^{G} \exp I_{g'}} \\
&= \frac{\sum_{k\in\mathcal{J}_g} \exp(\delta_k/(1-\sigma))^{1-\sigma}}{\sum_{g'=0}^{G}\left(\sum_{k\in\mathcal{J}_{g'}} \exp(\delta_k/(1-\sigma))\right)^{(1-\sigma)}}
\end{aligned}$$

Then the market share of product is:

$$\begin{aligned}
s_j &= s_{j|g}.\bar{s}_g \\
&= \frac{\exp\big(\delta_j/(1-\sigma)\big)}{\sum_{g'=0}^{G}\left(\sum_{k\in\mathcal{J}_{g'}} \exp(\delta_k/(1-\sigma))\right)^{1-\sigma}} \cdot \frac{1}{\sum_{k\in\mathcal{J}_g} \exp(\delta_k/(1-\sigma))^{\sigma}}
\end{aligned}$$

Nested logit model

Normalization of the outside good mean utility to 0:

$$s_0 = \frac{1}{1 + \sum_{g=1}^{G} \left(\sum_{k \in \mathcal{J}_g} \exp(\delta_k/(1-\sigma)) \right)^{1-\sigma}}$$

Market share inversion to obtain ξ (or equivalently, δ). We have:

$$\begin{aligned} & \ln(s_j/s_0) = \delta_j/(1-\sigma) - \sigma \ln \textstyle\sum_{k \in \mathcal{J}_g} \exp(\delta_k/(1-\sigma)) \\ \text{and} \quad & \ln s_{j|g} = \delta_j/(1-\sigma) - \ln \textstyle\sum_{k \in \mathcal{J}_g} \exp(\delta_k/(1-\sigma)) \end{aligned}$$

So that:

$$\ln s_j - \ln s_0 - \sigma \ln s_{j|g} = \delta_j$$

After rearranging we get:

$$\ln s_j - \ln s_0 = X_j\beta - \alpha p_j + \sigma \ln s_{j|g} + \xi_j$$

As before, we have a linear equation that we can estimate using 2sls. We have a new parameter σ and we need to instrument $\ln(\bar{s}_{j|g})$ too

Random coefficients model

As with micro data, we introduce unobserved individual heterogeneity

$$\begin{aligned} U_{ik} &= X_k \beta_i + \xi_k + \varepsilon_{ik} \\ &= \delta_k + \mu_{ik} + \varepsilon_{ik} \end{aligned}$$

Assume a (multivariate) distribution for $F_\beta(\bar{\beta}, \Sigma^X)$ so that:
$\beta_i = \bar{\beta} + \Sigma^X \nu_i$
$\delta_k = \bar{\beta} X_k + \xi_k$: "mean utility"
$\mu_{ik} = \Sigma^X \nu_i X_k$: "interaction term"

Two challenges: (i) no more closed-form solution for market shares; (ii) more complicated market share inversion

$$s_k\left(\delta; \Sigma^X\right) = \int \frac{\exp\left(\delta_k + \mu_k(\nu, \Sigma^X)\right)}{1 + \sum_{j=1}^{J} \exp\left(\delta_j + \mu_j(\nu, \Sigma^X)\right)} dF(\nu)$$

First challenge: use simulations for numerical approximation

Random coefficients model

Second challenge, how to invert market shares to recover $\xi(\theta)$?
Remark: once we have δ, we have ξ with a simple linear transformation

Berry and BLP prove that the market share system of equations can be inverted to find a unique vector of δ, for a given vector of parameters, i.e. they show that:

$$\delta_k = f^{-1}(s_0, s_1, ..., s_J, \Sigma^X)$$

Random coefficients model

Sketch of the proof that there is a unique $(\delta_1, ..., \delta_J)$ such that $s(\delta_1, ..., \delta_J) = s^{\text{obs}}$, which constitutes a system of non-linear equations

A sufficient condition is that the Jacobian of the matrix that represents the system of equations is diagonal-dominant
The Jacobian matrix is:

$$\begin{pmatrix} \frac{\partial s_1}{\partial \delta_1} & \cdots & \frac{\partial s_J}{\partial \delta_1} \\ \vdots & & \vdots \\ \frac{\partial s_1}{\partial \delta_J} & \cdots & \frac{\partial s_J}{\partial \delta_J} \end{pmatrix}$$

Diagonal-dominant means:

$$\left| \frac{\partial s_j}{\partial \delta_j} \right| > \sum_{k \neq j} \left| \frac{\partial s_k}{\partial \delta_j} \right|$$

Intuition: variation in δ_j must affect more s_j than the others products $k \neq j$ jointly

Random coefficients model

Trick:

$$\frac{\partial s_j}{\partial \delta_j} + \sum_{k \neq j} \frac{\partial s_k}{\partial \delta_j} + \frac{\partial s_0}{\partial \delta_j} = 0$$

Since $\frac{\partial s_j}{\partial \delta_j} > 0$, $\frac{\partial s_k}{\partial \delta_j}, \frac{\partial s_0}{\partial \delta_j} < 0$ We have: $\partial s_j / \partial \delta_j = |\partial s_j / \partial \delta_j|$ and $-\partial s_k / \partial \delta_j = |\partial s_k / \partial \delta_j|$ for $k \neq j$ or $k = 0$ So we have:

$$\left| \frac{\partial s_j}{\partial \delta_j} \right| = \sum_{k \neq j} \left| \frac{\partial s_k}{\partial \delta_j} \right| + \left| \frac{\partial s_0}{\partial \delta_j} \right|$$

So

$$\left| \frac{\partial s_j}{\partial \delta_j} \right| > \sum_{k \neq j} \left| \frac{\partial s_k}{\partial \delta_j} \right|$$

which is true as long $\left| \frac{\partial s_0}{\partial \delta_j} \right| > 0$, which comes from s_0>0: $\frac{\partial s_0}{\partial \delta_j} = -\int s_j(\nu) s_0(\nu) dF(\nu)$

Random coefficients model

We have established that if there exists a vector of δ such that $s(\delta, \Sigma^X) = s^{obs}$, it is unique.

How to solve for it?

By iteration, with a contraction mapping:

$$(\delta_k)^t = (\delta_k)^{t-1} + \log(s_k^{obs}) - \log(s_k^{theo}(\delta^{t-1}, \Sigma^X))$$

General proof that a function f is contracting:
(i) $f(\delta) = \delta$
(ii) $\exists \beta < 1$ s.t. $\forall x, x'$:

$$||f(x) - f(x')|| \leq \beta ||x - x'||$$

Sufficient conditions for (ii) to be satisfied are:

$$\frac{\partial f_j(\delta)}{\partial \delta_k} \geq 0 \qquad \forall k$$
$$\sum_k \frac{\partial f_j(\delta)}{\partial \delta_k} < 1$$

Estimation Method

Based on moment conditions:

- $\mathbb{E}\left[\xi Z^m\right] = 0$ for $m = 1, ..., M$ (no. of instruments)
- Use empirical counterparts of expectations = means over products and markets
- Compute the empirical counterparts of the moment conditions and use a weighting matrix to obtain the objective function to be minimized

How to recover ξ_k from δ_k?

- $\xi_k = \delta_k - X_k \bar{\beta}$
- It is also the residual of the IV regression of δ on X (IV because X includes the price)
- So the linear parameters $\bar{\beta}$ are a deterministic function on the non-linear parameters, so we can "integrate out" the linear parameters from the objective function
- It is very useful since it implies that the dimension of optimization does not increase when we introduce brand, product, time fixed effects

Substitution patterns

$$\frac{\partial s_j}{\partial p_k} = \int -\alpha s_j(\alpha, \beta) s_k(\alpha, \beta) dF(\alpha, \beta)$$

Assume products j and k have both a high value of characteristic X^l. Individual with a high β_l have high market shares for both products with high X^l, so large substitution between the two products; while for individuals with low β^l, low substitution

$$\frac{\partial s_j}{\partial p_j} = \int \alpha s_j(\alpha, \beta)(1 - s_j(\alpha, \beta)) dF(\alpha, \beta)$$

Demand for cars

- Berry, Levinsohn and Pakes, 1995
  - U.S. automobile market
  - Many distinct car models
  - Car model are described by characteristics
  - Typically buy only one car
- Goal: Reasonable cross price elasticity estimates
  - Based on discrete choice model
- Aggregate market data
  - Estimate model based on market shares, prices and product characteristics

Demand specification

Utility model with random coefficients

$$
\begin{aligned}
U_{ik} &= \alpha \ln (y_i - p_k) + x_k' \beta + \xi_k + \sum_l \sigma_l x_{kl}' \nu_{il} + \varepsilon_{ik} \\
&= \delta_k + \mu_{ik} + \varepsilon_{ik}
\end{aligned}
$$

where ν_{il} are drawn from a known distribution function, and measure a consumer's preference for specific attributes (such as size of the car).

Utility specification accounts for:

- Consumer and product characteristics interaction
- Income distribution

As before prices may be endogenous due to unobserved characteristics ξ_k (style, dealer quality etc.)

Instruments

Instruments:

- Cost shifters
- Characteristics of other products (functions, such as sum or average)

Because of competition, characteristics of other products affect prices (by constraining market power): $\mathbb{E}(Z|p) \neq 0$

Characteristics of other products are not correlated with demand shock ξ

Here, utility of product k depends only on characteristics of product k, and does not depend on the characteristics of other products

Data

- Automotive News Market Data Book

  Describes model characteristics (weight, horse power, Air conditioning, Miles per Gallon, Size, etc.), prices in 1983 dollars of base models, US sales by name plate
- Annual data, 1971-1990
- Additional data sources for income distribution, etc.

Results

Logit model (selected estimates):

	OLS - Logit	IV - Logit
Constant	-10.07	-9.27
	(0.25)	(0.49)
HP/Weight	-0.12	1.97
	(0.28)	(0.91)
Size	2.34	2.36
	(0.13)	(0.25)
Price	-0.09	-0.22
	(0.00)	(0.12)

- Price coefficient estimates are low
- Price endogenous
- Estimates imply inelastic demand curves (elasticities between 0 and -1)

Results

BLP Model (selected demand estimates):

	Parameter	Std
Constant	-7.06	(0.49)
HP/Weight	2.88	(2.02)
Size	3.46	(0.61)
$ln(y - p)$ (α-coefficient)	43.50	(6.43)
(...)		

Results

- Price coefficient estimates are high in contrast to logit estimates
- Estimates imply elastic demand curves
- Price elasticity estimates range between -3 and -120 for most models
- Semi cross-price elasticities (% market share change of row car with a $1,000 price increase in the column car)

	Accord	Ford Taurus	Mazda 323
Accord	-51.64	1.53	0.12
Taurus	2.04	-43.63	0.06
323	2.18	0.85	-125.93

Presentation

Nevo (2000)

- Ready-to-Eat Cereal Industry:
  - Many products
  - Few firms
  - High price cost margins
- Goal:
  - Reasonable cross price elasticity estimates
  - Counter-factual simulation: horizontal merger
  - Post-merger equilibrium
  - Welfare analysis
- Aggregate market data

Model Specification

Utility of product j in market t to consumer i

$$U_{ijt} = x_j'\beta_i - \alpha_i p_{jt} + \xi_j + \triangle\, \xi_{jt} + \varepsilon_{ijt}$$

where:

- x are observed product characteristics such as calories sodium, and fiber content
- p is price
- ξ_j is unobserved product characteristic
- $\triangle\, \xi_{jt}$ is city-quarter deviation from the mean unobserved product characteristic
- ε is a mean zero error term

Unobserved heterogeneity:

$$\begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \Pi \mathbf{D}_i + \Sigma \mathbf{v_i} \quad \text{with} \quad v_i \sim N(\mathbf{0}, \mathbf{I})$$

where D are demographic variables, Π is a coefficient matrix and Σ is the standard-deviation matrix

Demand-side

Recover demand shocks ξ_{jt} as before and interact with instruments

As before, prices may be endogenous due to unobserved characteristics

Instruments: Average price of a brand in all other cities

Valid under the assumption is that city specific valuations are independent across cities but possibly correlated within a city

Supply-side

Supply side as before:

$$\pi_f = \sum_{j \in J_f} \left[p_{jt} - mc_{jt}\right] \cdot M \cdot s_{jt}\left(p, x, \xi; \alpha, \beta\right)$$

FOC:

$$s_{jt}\left(p, x, \xi; \alpha, \beta\right) - \sum_{l \in J_f} \left[p_{lt} - mc_{lt}\right] \frac{\partial s_{lt}\left(p, x, \xi; \alpha, \beta\right)}{\partial p_{kt}} = 0$$

Recover marginal cost shock u_{jt}:

$$\begin{aligned} mc_{jt} &= w_{jt}\gamma + u_{jt} \\ p_{jt} - b(p, x, \xi; \alpha, \beta) &= w_{jt}\gamma + u_{jt} \end{aligned}$$

Estimation method

GMM:

$$\omega(\theta) = (\xi(\theta), u(\theta))$$

Let Z denote instruments such that

$$\mathbb{E}[Z'\omega(\theta)] = 0$$

Estimator $\omega(\theta)$ solves:

$$\min_{\theta} \omega(\theta)' Z A^{-1} Z' \omega(\theta)$$

for a suitable weight matrix A.

Data

- UConn Infoscan data base.
- Time period: Q1 1988 - Q4 1992.
- 65 US cities.
- Prices and market shares in city-quarter pairs, brand characteristics, advertising and demographic information.

Results

Selected demand estimates:

	Parameter	Std
Price	-27.198	(5.25)
Advertising	0.020	(0.01)
Calories from Fat	1.146	(0.13)
Sugar	5.742	(0.58)
(...)		

- Price coefficient estimates are high.
- Estimates imply elastic demand curves.
- Price elasticity estimates range between -2 and -5 for most cereals.

Results

- Cross price elasticities:

  % change in market share of brand *i* with a one percent change in price of *j* (*i* row, *j* column)

	Corn Flakes	Frosted Flakes	Rice Krispies
K Corn Flakes	-3.38	0.21	0.20
K Raisin Brand	0.04	0.05	0.08
K Frosted Flakes	0.15	-3.14	0.11

Results

Price cost margins: (p-mc)/p

Observed and simulated for different market structure (potential mergers)

	Median	Low (2.5%)	High (97.5%
Single Product Firm	35.8%	24.4%	46.4%
Curr. ownership of 25 Brands	42.2%	29.1%	55.8%
Joint ownership of All Brands	72.6%	62.2%	97.2%

Limitations

Limitations of the BLP type setting:

- Static demand: Cars are durable goods
- Replacement decisions are ignored
- Technological changes may induce structural changes in parameters (car industry experienced substantial changes during 1970s and 1980s)

General concerns:

- Aggregate demand model based on functional form assumptions
- May be adequate for substitution patterns
- May not explain consumers' preferences
- Each consumer buys at most one good (discrete choice)