Demand for differentiated products

TSE, MRes

Outline

Theory and estimation with aggregate data

To go further:

- What to do with micro data?
- With a combination of aggregate and micro data

Demand for differentiated products

Products are differentiated

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

Consequence: 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 (marketing and pricing departments)

From demand and prices, we can infer margins, mark up which indicate firms' market power (see Lecture 2)

Understanding product demand is crucial for sectoral policies (e.g carbon tax)

Outline

Reduced form and 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

Presentation of the BLP model

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

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

With product-level aggregate data, the BLP model is able to:

- Include product differentiation through observed and unobserved product characteristics
- Control for price endogeneity
- Include unobserved individual heterogeneity in preferences

Very influential model, used a lot by researchers and practitioners (competition authorities)

Remark: BLP also develop a supply model, see next lecture.

Overview of the model

Discrete choice model, choice of one product among the set of products available

One option is not to purchase (outside option), typically denoted by 0

Model is micro-founded: starts from specification of utility, as a function of characteristics of products (including price)

Individuals do not have preference for products but for product characteristics

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

Each consumer chooses the option associated to the highest utility

Basic idea: aggregate sales represent the aggregation of individual optimal choices

Objective = estimate parameters of utility from aggregate sales data

The BLP model

$$U_{ik} = \mathbf{x}_k' \boldsymbol{\beta}_i - \alpha_i p_k + \xi_k + \varepsilon_{ik}$$

- $\mathbf{x}_k'$: vector of observed product characteristics
- p_k: price of the product, endogenous: $\mathbb{E}(\boldsymbol{\xi}|\mathbf{p}) \neq 0$
- ξ_k: unobserved product characteristics (unobserved to the econometrician, but known by individuals and firms)
- α_i price sensitivity
- $\boldsymbol{\beta}_i$ vector of preference for product characteristics
- ε_{ik}: individual and product taste shock, iid across individuals and products

Important assumption: ξ_k are the same for all consumers, no unobserved heterogeneity in the valuation of the unobserved characteristics (see Berry & Haile, 2014)

The BLP model

We typically assume $\alpha_i, \boldsymbol{\beta}_i$ are drawn from known distributions. And we want to estimate the parameters of the distributions. We can usually decompose the individual parameters into a mean and an individual deviation:

$$\alpha_i = \bar{\alpha} + \sigma^p \zeta_i^p$$
$$\beta_i^l = \bar{\beta}^l + \sigma^l \zeta_i^l$$

rk: l index for the characteristics.

$\bar{\alpha}$ and $\bar{\beta}^l$ are the mean parameters and σ^p and σ^l are the standard deviations. ζ_i^p and ζ_i^l are draws from a known distribution.

We can rewrite the utility function as:

$$U_{ik} = \delta_k + \mu_{ik} + \varepsilon_{ik}$$

- δ_k: mean product utility common to everyone: $\delta_k = \sum_l x_k^l \bar{\beta}^l - \bar{\alpha} p_k + \xi_k$
- μ_{ik}: individual-specific deviation from the mean utility $\mu_{ik} = \sum_l x_k^l \sigma^l \zeta_i^l - p_k \sigma^p \zeta_i^p$

The logit model

The error terms ε_{ik} are assumed to be extreme value distributed across individuals and products

Cumulative distribution function:

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

Probability 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 when 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: $t = \exp(-x), dt = -\exp(-x)dx$

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

The simple logit model

We first examine two restricted versions of the BLP model: logit and nested logit models

The simple logit model assumes homogeneous preferences: $\alpha_i = \alpha$, $\boldsymbol{\beta}_i = \boldsymbol{\beta}\ \forall i$

Individuals only differ in their product-specific tastes ε_{ij}

Model rules out systematic individual preferences towards some product characteristics (e.g. preference for small cars rather than large ones)

$$s_{ik} = s_k(\alpha, \beta) = \frac{\exp\left(\mathbf{x}_\mathbf{k}'\boldsymbol{\beta} - \alpha p_k + \xi_k\right)}{\sum_{j=0}^{J} \exp\left(\mathbf{x}_\mathbf{j}'\boldsymbol{\beta} - \alpha p_j + \xi_j\right)}$$

Estimation of the simple logit model

Idea of estimation: match observed market shares s_k^{obs} with those predicted by the model $s_k(\alpha, \boldsymbol{\beta})$

However: $s_k(\alpha, \boldsymbol{\beta})$ depend on the vector of unobserved characteristics $\boldsymbol{\xi}$

Idea: consider $\boldsymbol{\xi}$ to be the error term which distribution is not specified

Estimation using moment conditions (GMM) from orthogonality condition:

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

So we need to express $\boldsymbol{\xi}$ as function of observed market shares, parameters and exogenous variables X: $\boldsymbol{\xi}(\mathbf{s}^{\text{obs}}, \mathrm{X}, \alpha, \boldsymbol{\beta})$

We need a “market share inversion” procedure

Estimation of the simple logit model

Trick: express the share of the "outside good", which mean utility δ_0 is normalized to 0

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

Implies:

$$\begin{aligned} \log\left(s_k\right) - \log\left(s_0\right) &= \delta_k \\ &= \mathbf{x}_{\mathbf{k}}'\boldsymbol{\beta} - \alpha p_k + \xi_k \end{aligned}$$

So we can easily invert the market share equations and get:

$$\xi_k(\alpha, \boldsymbol{\beta}) = \log\left(s_k\right) - \log\left(s_0\right) - \mathbf{x}_{\mathbf{k}}'\boldsymbol{\beta} + \alpha p_k$$

Estimation of the simple logit model

We have a linear equation to estimate:

$$\underbrace{\log(s_k) - \log(s_0)}_{\text{dependent variable}} = \underbrace{\mathbf{x}_\mathbf{k}'}_{\text{exogenous}} \boldsymbol{\beta} - \alpha \underbrace{p_k}_{\text{endogenous}} + \underbrace{\xi_k}_{\text{residual}}$$

GMM is equivalent to an IV regression, we can use 2SLS to estimate $\boldsymbol{\beta}$ and α

We need instruments for price (at least one):

- correlated with price $\mathbb{E}(\mathbf{p}|\mathbf{Z}) \neq 0$
- uncorrelated with ξ_k the demand shock or unobserved product characteristics $\mathbb{E}(\boldsymbol{\xi}|\mathbf{Z}) = 0$

Instruments

Traditional instruments:

- Cost shifters (rk: need cost shifters that vary across products, not very common)
- Functions of other product characteristics, “BLP instruments”: instrument p_k by $\sum_{j \neq k} x_j^l$
- Intuition = a product’s margin and price are 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, “Hausman instruments” (e.g. different cities)
- Prices in different cities correlated through the costs but independent of the local demand shock
- Instruments invalid if there are common demand shocks that affect prices (e.g. national advertising campaign)

Further discussion on instruments later.

Limitations of the simple logit

Price elasticities (rk: $\alpha > 0$):

$$\begin{aligned}\frac{\partial s_k}{\partial p_k} &= -\alpha \frac{\partial s_k}{\partial \delta_k} = -\alpha s_k(1 - s_k) \\ \frac{\partial s_k}{\partial p_j} &= \alpha s_k s_j\end{aligned}$$

Substitution between two products only depend on market shares

We would prefer substitution to depend on characteristics of products: more substitution within products with similar characteristics

IIA (independence from irrelevant alternatives): the ratio of market shares of two products does not depend on other products

$$\frac{s_j}{s_k} = \frac{\exp\left(X_j\beta - \alpha p_j + \xi_j\right)}{\exp\left(X_k\beta - \alpha p_k + \xi_k\right)}$$

Solutions:

- Nested Logit model
- Random coefficients model

Nested Logit Model

Idea: Group products that are similar and assume a 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

Remarks:

- 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 groups 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

$$U_{ij} = \mathbf{x}_j'\beta - \alpha_i p_j + \xi_j + \zeta_{ig} + (1-\sigma)\epsilon_{ij}$$

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

Within-nest market shares: ζ_{ig} is irrelevant because it is common to all products in the nest

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

Same as logit probability, except δ replaced by $\delta/(1-\sigma)$

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

Nested Logit Model

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

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

Probability of choosing a nest $\bar{s}_g$:

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

So the product market share is:

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

Remark: nest 0 contains only the outside good

Nested logit model

Normalization of the outside good mean utility to 0:

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

We use the same trick for the market share inversion, i.e. take the logarithm of s_j/s_0.

We have:

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

So that:

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

Nested logit model

After rearranging we get:

$$\log s_j - \log s_0 = \mathbf{x}_j'\beta - \alpha p_j + \sigma \log s_{j|g} + \xi_j$$

As before, we have a linear equation that we can estimate using 2sls.

We have a new parameter (σ) to estimate and we need to instrument $\log(\bar{s}_{j|g})$

Extension of BLP instruments: sum of product characteristics within nests: instrument $\log(\bar{s}_{j|g})$ by $\sum_{k\neq j, k\in \mathcal{G}_j} X_k$, where $\mathcal{G}_j$ denotes the nest of product j

Random coefficients model

We introduce unobserved individual heterogeneity

$$\begin{aligned} U_{ik} &= \mathbf{x}_k'\boldsymbol{\beta}_i - \alpha_i p_k + \xi_k + \varepsilon_{ik} \\ &= \delta_k + \mu_{ik} + \varepsilon_{ik} \end{aligned}$$

Assume a (multivariate) distribution for the parameters of preferences $F_{\beta}(\bar{\boldsymbol{\beta}}, \boldsymbol{\Sigma}^x)$ so that: $\beta_i^l = \bar{\beta}^l + \sigma_l \nu_i^{x,l}$
$\alpha_i = \bar{\alpha} + \sigma^p \nu_i^p$
$\delta_k = \mathbf{x}_k'\bar{\boldsymbol{\beta}} + \xi_k$: "mean utility"
$\mu_{ik} = \mathbf{x}_k'\boldsymbol{\Sigma}^x \boldsymbol{\nu_i}^x - \sigma^p \nu_i^p p_k$: "interaction term"

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

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

Market share inversion

First challenge: use numerical simulations for an approximation of the integral. We draw vectors of $\boldsymbol{\nu}$ in the (possibly joint) distribution $F_{\boldsymbol{\nu}}(.)$, compute the expression inside the integral for each draw, and take the average.

Second challenge, how to invert market shares to recover $\boldsymbol{\xi}(\theta)$?

Preliminary remark: once we have $\boldsymbol{\delta}$, we have $\boldsymbol{\xi}$ with a simple linear transformation:

$$\xi_k = \delta_k - \mathbf{x}'_k \bar{\boldsymbol{\beta}}$$

Berry (1994) and BLP (1995) prove that the market share system of equations can be inverted to find a unique vector of $\boldsymbol{\delta}$, for a given vector of parameters, i.e. they show that:

$$\delta_k = f^{-1}(\mathbf{s}^{\text{obs}}, \boldsymbol{\Sigma}^x, \sigma^p, \mathrm{X}, \mathbf{p}, F_{\boldsymbol{\nu}})$$

is unique.

Market share inversion

Sketch of the proof that there is a unique $\delta = (\delta_1, \ldots, \delta_J)$ such that $\mathbf{s}(\delta_1, \ldots, \delta_J) = \mathbf{s}^{\text{obs}}$

This constitutes a system of non-linear equations in δ

A sufficient condition for unicity is that the Jacobian of the matrix that represents the system of equations is diagonal-dominant (see Gale & Nikaidô, 1965)

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:

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

Market share inversion

Intuition: variation in δ_j must affect more s_j than the shares of all the others products $k \neq j$, which is ensured by the existence of the outside option

Formal proof. We have:

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

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:

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

Market share inversion

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 as soon as $\left|\frac{\partial s_0}{\partial \delta_j}\right| > 0$, which comes from $s_0 > 0$:

$$\frac{\partial s_0}{\partial \delta_j} = -\int s_j(\boldsymbol{\nu}) s_0(\boldsymbol{\nu}) dF_{\boldsymbol{\nu}}(\boldsymbol{\nu})$$

The full proof is in Berry 1994

We have established uniqueness of δ, now the question is how to find its expression since there is no closed form solution?

Random coefficients model

We have established that if there exists a vector of δ such that $s(\boldsymbol{\delta}, \boldsymbol{\Sigma}^x, \sigma^p, \mathrm{X}, \mathbf{p}, F_{\boldsymbol{\nu}}(\boldsymbol{\nu})) = s^{obs}$, it is unique.

The question is now how to solve for it?

By iteration, with BLP contraction mapping:

$$(\delta_k)^t = (\delta_k)^{t-1} + \log(s_k^{obs}) - \log(s_k^{theo}\left(\boldsymbol{\delta}^{t-1}, \boldsymbol{\Sigma}^x, \sigma^p, \mathrm{X}, \mathbf{p}, F_{\boldsymbol{\nu}}(\boldsymbol{\nu})\right)$$

General proof that a function f is contracting:
(i) $f(\boldsymbol{\delta}) = \boldsymbol{\delta}$
(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:

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

Full proof in BLP (1995)

Estimation Method

Based on moment conditions:

- $\mathbb{E}\left[\boldsymbol{\xi}\mathbf{z}^m\right] = 0$ 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?

- $\xi_k = \delta_k - \mathbf{x}_k'\bar{\boldsymbol{\beta}} - \bar{\alpha}p_k$
- It is also the residual of the IV regression of $\boldsymbol{\delta}$ on $[\mathrm{X}, \mathbf{p}]$ (IV because of the price)
- So the linear parameters $\bar{\boldsymbol{\beta}}$ 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 a lot of controls (product characteristics, brand, product, time fixed effects)

Substitution patterns

$$\frac{\partial s_j}{\partial p_k} = \int -\alpha s_j(\boldsymbol{\nu}) s_k(\boldsymbol{\nu}) dF_{\boldsymbol{\nu}}(\boldsymbol{\nu})$$

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

$$\frac{\partial s_j}{\partial p_j} = \int (\bar{\alpha} + \sigma^p \nu^p) s_j(\boldsymbol{\nu})(1 - s_j(\boldsymbol{\nu})) dF(\boldsymbol{\nu})$$

Identification

Which variation in the data identifies the parameters of preferences?

Common utility parameters: correlation between choices and product characteristics (across products variation)

Parameters of the unobserved heterogeneity: identification leverages heterogeneity in choices under different choice sets

Thought experiment: two markets (periods), in one market a product is not available, substitution towards other product should be informative about heterogeneity in preferences for a certain characteristics, if everyone substitutes to the same product –> homogeneous preferences

Advances on instruments

"Recent" instruments:

- Gandhi & Houde (WP, latest version 2/9/2020!): differentiation instruments
- Unpublished paper but instruments have been used a lot recently
- Idea: what drives demand for product j is not the level of its characteristics x_j^l but its difference with competing products
- Use flexible approximations of an unknown function of the difference (quadratic and spline)
- Quadratic diff IV: $\sum_{j'}(d^l_{jt,j'})^2, \quad \sum_{j'}(d^{\hat{p}}_{jt,j'})^2$
  - where $d^l_{j,j'} = x^l_{j'} - x^l_j$
  - $d^{\hat{p}}_{j,j'} = \hat{p}_{j'} - \hat{p}_j$
  - $\hat{p}_j$ are the predicted values from a linear regression on instruments (1st stage)
- Local diff IV: $\sum_{j'} |d^l_{j,j'}| \mathbb{1}\left(|d^l_{j,j'}| < \kappa^l\right), \quad \sum_{j'} |d^{\hat{p}}_{j,j'}| \mathbb{1}\left(|d^{\hat{p}}_{j,j'}| < \kappa^{\hat{p}}\right)$
  - κ are a proximity thresholds (e.g. standard-deviations)

Instruments

"Recent" instruments (cont'd):

- Approximation of Chamberlain optimal instruments (Reynaert & Verboven, 2014):
  - Chamberlain (1987) optimal instruments: $\mathbb{E}\left(\frac{\partial \xi_k(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} | \mathbf{Z}\right)$
  - Need approximations
  - 1) Approximated by $\mathbb{E}\left(\frac{\partial \xi_k(\hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}} | \mathbf{Z}\right)$, need a first estimate of $\boldsymbol{\theta}$, with standard instruments
  - 2) Evaluate the functions at the expected value of the unobservable $\mathbb{E}(\boldsymbol{\xi}) = 0$ to avoid integrating over the distribution of $\boldsymbol{\xi}$
- Exogenous policy shocks (e.g. tax change)
- City-time dummies, to take advantage of control and treated groups in the spirit of DiD (see Li, 2017 Restud)