Lecture 3: Price discrimination

TSE, MRes

Outline

1 Introduction

2 Unobserved price discrimination

Definitions

Three types of price discrimination:

- First-degree, or perfect price discrimination: each individual i pays a different price p_i that reflects its willingness to pay. In this case and under monopoly, all the consumer surplus is extracted by the firm. Under imperfect competition, not clear because competition for each consumer intensifies
- Second degree: firms offer menus of contract such that consumers self select. Typically, firms produce different versions of the product with different quality levels associated with different prices (p_v, q_v)
- Third degree: firms observe some consumers characteristics that are correlated with their preferences and set prices conditional on these observables. Consumers are segmented in $d = 1, ..., n_D$ groups and firms set p_d. First-degree PD is a limit case of third-degree PD. Welfare impact of price discrimination unclear; typically winners and losers but depends on competition

Second degree price discrimination

Most firms offer more than one version of their products (e.g. computer, cars, education,...)

Each version of the product associated with different characteristics and a different price

Second degree price discrimination when the price difference between two versions is not equal to the cost difference which amounts to test:

$$p_{jv} - c_{jv} = p_{jv'} - c_{jv'}$$

Third degree price discrimination

A lot of examples

Third degree PD can be explicit: half price for women, reduced-price for kids and seniors

Or implicit: different prices in different cities, different prices for train or flight tickets depending on the purchase date, new consumer discounts, unobserved rebates...

But, the extent of third degree PD is limited by arbitrage opportunities (e.g driving to a cheaper city or resale across consumer groups...)

Questions of interest

Detecting price discrimination: are price differences attributable to cost or markup differences?

Welfare impacts of price discrimination: who benefits from price discrimination? Would uniform pricing be welfare-improving?

Heterogeneous policy effects very relevant since different prices are paid by different groups of consumers (e.g. heterogeneous impacts of a merger, a tax...)

Outline

1. Introduction
2. Unobserved price discrimination

Motivation

Based on D'Haultfœuille, Durrmeyer & Février (Restud, 2019)

Problem of limited data on prices paid by consumers when estimating demand and supply for differentiated product markets

Transaction prices difficult to collect, typically we can observe

- Posted prices, MSRP, maximum price
- Average prices
- Individual prices for products purchased, but not for products that the individual ended up not buying

We propose a method to deal with limited data on prices, when price dispersion is due to **price discrimination**

In the model, consumers pay different (unobserved) prices because they have different preferences

Motivation

Why do we care?

- Difference between transaction prices and observed price is a non-classical measurement error
- We expect transaction prices not to be assigned “randomly” to consumers
- Standard methods fail to estimate consistently demand and supply under unobserved price discrimination
- Crucial to know the transaction prices for welfare analysis

Overview of the method

By putting structure on how transaction prices are determined, we are able to estimate demand and supply parameters, together with unobserved transaction prices

Main idea: replace unobserved prices by the prices implied by the FOC associated to profit maximization

We rely on two crucial assumptions:

- There is a known mapping between the observed prices and the marginal costs of products
- Marginal cost is independent of the buyer's identity

Overview of the method

Intuition of the approach:

- For a given vector of parameters of preferences, we back out marginal costs from the observed prices
- Using the marginal cost we can express all the unobserved transaction prices using the FOC
- We apply the standard BLP method to estimate demand except that we use "theoretical" transaction prices when they are unobserved

Simple framework

To illustrate the method, look at the model with simple logit demand (no random coefficient) and single product firms

Introduce a new index d for demographic group, i.e. group of consumers who pay the same price. We specify the utility function:

$$
\begin{aligned}
U_{ij}^{d} &= \mathbf{x}_j' \boldsymbol{\beta}^d + \alpha^d p_j^d + \xi_j^d + \epsilon_{ij}^d \\
&= \delta_j^d + \epsilon_{ij}^d
\end{aligned}
$$

- $\mathbf{x}_j$ observed product characteristics
- ξ_j^d unobserved characteristics (to the econometrician, they are observed by both consumers and firms)
- p_j^d discriminatory price (group specific)
- ϵ_{ij}^d, iid error term (Extreme Value)
- $\boldsymbol{\theta}^d = (\boldsymbol{\beta}^d, \alpha^d)$ to estimate

The simple logit case

Aggregation of individual choices within a demographic group:

$$s_j^d = \frac{\exp\left(\delta_j^d\right)}{\sum_{k=0}^{J} \exp\left(\delta_k^d\right)}$$

Usual normalization: $\delta_0^d = 0$ so that $\ln\left(s_j^d / s_0^d\right) = \delta_j^d$

We cannot directly obtain the residual as:

$$\xi^d = \ln\left(s_j^d / s_0^d\right) - \mathbf{x}_j \boldsymbol{\beta}^d - \alpha^d p_j^d$$

because p_j^d are unobserved. Instead, we specify p_j^d from the supply-side

Note that replacing p_j^d by $\bar{p}_j$ the observed price leads to bias since $\bar{p}_j - p_j^d$ are likely to be correlated with product characteristics which makes the standard BLP instruments invalid

Simple supply-side model

Nash-Bertrand equilibrium with single-product firms and price discrimination

Consumers segmented into n_D groups. Profit for firm f producing good j:

$$\Pi_f = \sum_{d=1}^{n_D} N_d s_j^d \times \left(p_j^d - c_j\right)$$

Optimal discriminatory price:

$$p_j^d = c_j - \frac{s_j^d}{\partial s_j^d / \partial p_j^d}$$

Under the logit demand, we have: $\partial s_j^d / \partial p_j^d = \alpha^d s_j^d (1 - s_j^d)$ so that

$$p_j^d = c_j - \frac{1}{\alpha^d (1 - s_j^d)}$$

Identification

For each product, find the **pivot group** as consumer group with highest price (=highest mark-up):

$$\bar{d}_j = \arg \max_{d=1,\dots,n_D} \frac{-1}{\alpha^d (1 - s_j^d)}$$

From FOC for pivot group, recover marginal cost $c_j = \overline{p}_j + \frac{1}{\alpha^{\bar{d}_j}(1-s_j^{\bar{d}_j})}$

Unobserved prices p_j^d such that:

$$p_j^d = \overline{p}_j - \frac{1}{\alpha^d (1 - s_j^d)} + \frac{1}{\alpha^{\bar{d}_j}(1 - s_j^{\bar{d}_j})}$$

Get the residuals:

$$\xi_j^d = \ln\left(s_j^d / s_0^d\right) - \mathbf{x}_j' \boldsymbol{\beta}^d - \alpha^d \bar{p}_j - \frac{1}{(1 - s_j^d)} + \frac{\alpha^d}{\alpha^{\bar{d}_j}(1 - s_j^{\bar{d}_j})}$$

Identification and estimation

Construct moment conditions:

$$\mathbb{E}\left[\mathbf{z}_j^{d\prime}\xi_j^d(\alpha^1, \ldots, \alpha^{n_D})\right] = 0$$

Estimation by GMM since $\boldsymbol{\xi}^d$ are a non-linear function of $\alpha^1, \ldots, \alpha^{n_D}$

Extension to multiproduct firms straightforward

Extension to RC model without random coefficient on the price OK

Extension to RC model with random coefficient on price more involved

General BLP demand model

Utility function:

$$U_{ij}^{d} \quad = \mathbf{x}_j' \boldsymbol{\beta}_i^{d} + \alpha_i^{d} p_j^{d} + \xi_j^{d} + \epsilon_{ij}^{d}$$

Heterogeneity of preferences:

$$\begin{pmatrix} \beta_{il}^{d} \\ \alpha_i^{d} \end{pmatrix} = \begin{pmatrix} \bar{\beta}_{x_l}^{d} + \sigma_{x_l}^{d} \nu_i^{x_l} \\ \bar{\alpha}^{d} + \sigma_p^{d} \nu_i^{p} \end{pmatrix}$$

Market share equation for demographic group d:

$$s_j^{d} = \int s_j^{d}(\boldsymbol{\delta}^{d}, \mathbf{x}, \mathbf{p}^{d}, \boldsymbol{\nu}^{d}) dF(\boldsymbol{\nu}) = \int \frac{\exp\left(\delta_j^{d} + \mu(\boldsymbol{\nu}^{d}, \mathbf{x}_j, p_j^{d})\right)}{\sum_{k=0}^{J} \exp\left(\delta_k^{d} + \mu(\boldsymbol{\nu}^{d}, \mathbf{x}_k, p_k^{d})\right)} dF(\boldsymbol{\nu}^{d})$$

Where $\mu(\boldsymbol{\nu}^{d}, \mathbf{x}_j, p_j^{d}) = \sum_l \sigma_{x_l}^{d} \nu^{x_l} x_j^{l} + \sigma_p^{d} \nu^{p} p_j^{d}$

General supply model

Profit for firm f producing the set of goods $\mathcal{J}_f$:

$$\Pi_f = \sum_{d=1}^{n_D} N_d \sum_{j \in \mathcal{J}_f} s_j^d \times \left(p_j^d - c_j\right)$$

Optimal discriminatory price:

$$p_j^d = c_j - \left[(\Omega^d)^{-1} S^d\right]_j$$

Where Ω_{jk}^d is the matrix of derivatives $\frac{\partial s_k^d}{\partial p_j^d}(\boldsymbol{\delta}^d, \boldsymbol{\nu}^d, \mathbf{x}, \mathbf{p}^d)$ multiplied by the ownership matrix, with:

$$\frac{\partial s_k^d}{\partial p_j^d}(.) = -\int (\alpha^d + \sigma^p \nu^p) s_j(\boldsymbol{\delta}^d, \boldsymbol{\nu}^d, \mathbf{x}, \mathbf{p}^d) s_k(\boldsymbol{\delta}^d, \boldsymbol{\nu}^d, \mathbf{x}, \mathbf{p}^d) dF(\boldsymbol{\nu}^d)$$
$$\frac{\partial s_j^d}{\partial p_j^d}(.) = \int (\alpha^d + \sigma^p \nu^p) s_j(\boldsymbol{\delta}^d, \boldsymbol{\nu}^d, \mathbf{x}, \mathbf{p}^d) \left(1 - s_j(\boldsymbol{\delta}^d, \boldsymbol{\nu}^d, \mathbf{x}, \mathbf{p}^d)\right) dF(\boldsymbol{\nu}^d)$$

If c_j and δ^d are known, we can recover p_j^d associated to a vector of parameters

Identification

δ^d not observed but solved from the inversion of market share

Issue: from BLP we know we can invert $\mathbf{s}^d$ in $\boldsymbol{\delta}^d$ if $\mathbf{x}$ and $\mathbf{p}^d$ are observed but p^d is not observed in our case.

Need to solve jointly for $(\boldsymbol{\delta}^d, \mathbf{p}^d)$ such that:

$$\begin{cases} \mathbf{s}^d(\boldsymbol{\delta}^d, \boldsymbol{\nu}^d, \mathbf{x}, \mathbf{p}^d) = \mathbf{s}^{d,obs} \\ \mathbf{p}^d = \mathbf{c} - (\Omega^d(\boldsymbol{\delta}^d, \boldsymbol{\nu}^d, \mathbf{x}, \mathbf{p}^d))^{-1}\mathbf{s}^{d,obs} \end{cases}$$

We propose an algorithm to solve to $\mathbf{p}^d$ and δ^d which extends BLP's contraction mapping and show that our algorithm is a contraction under a condition on σ_p^d

The estimation algorithm

Two inner loops algorithm, (outer loop = search over parameters to estimate): for each vector of parameters $\theta = (\sigma_x^d, \sigma_p^d)$

1. For initial values of transaction prices $(\mathbf{p}^{d,0})$ compute $(\boldsymbol{\delta}^d(\mathbf{p}^{d,0}))$ using BLP's contraction mapping
2. Determine pivot groups as:

$$\arg \max_{d=1,\ldots,n_D} -(\Omega^d(\boldsymbol{\delta}^d(\mathbf{p}^{d,0}), \boldsymbol{\nu}^d, \mathbf{x}, \mathbf{p}^{d,0}))^{-1}\mathbf{s}^{d,obs}$$

3. Estimate the marginal cost from $\bar{\mathbf{p}}$ and the FOC for pivot group and compute the transaction prices for the non-pivot groups
4. Update the values of transaction price $(\mathbf{p}^{d,1})$ and iterate until convergence of transaction prices
5. When convergence in transaction prices is achieved, get the residuals $\boldsymbol{\xi}^d$ and compute the GMM objective condition

The estimation algorithm

Note that the estimation algorithm is more computationally intensive since we need to invert the market share equation and solve for δ^d for every value of transaction prices until convergence of the *price-loop*

Rk: the model and algorithm are much simpler when there is no random coefficient on the price (but we can have random coefficients on observed characteristics $\mathbf{x}$)

Indeed, in this case we can solve separately for δ^d and $\mathbf{p}^d$ such that:

$$\begin{cases} \mathbf{s}^d(\boldsymbol{\delta}^d, \boldsymbol{\nu}^d, \mathbf{x}) = \mathbf{s}^{d,obs} \\ \mathbf{p}^d = \mathbf{c} - \left(\Omega^d(\boldsymbol{\delta}^d, \boldsymbol{\nu}^d, \mathbf{x})\right)^{-1}\mathbf{s}^{d,obs} \end{cases}$$

Application

New automobile market in France

- We observe prices from catalogs (MSRP), interpreted as maximal prices
- Evidence that discounts exist
- Evidence that some people pay the posted prices
- Major part of the marginal cost related to the production, not the sale

We consider price discrimination based on age and expected income (median income of town by age class):

- Easily observable by sellers
- Presumably strong determinants of purchases

Other applications, extensions

- Coupons, loyalty rebates
- Airline industry
- Movie theater
- International price dicrimination (e.g. drugs, cars, beer)
- Academic vs. business pricing

Extensions:

- Unobserved market share by demographic group, need additional restriction: $\xi_j^d = \xi_j, \forall d$
- Different costs across groups and cost differences related to observed cost shifters that are different from demand variables
- We do not observe the exact individual characteristic used for discrimination but a proxy
- Different supply models (vertical relation, collusion...)