Demand and supply of differentiated products

Paul Schrimpf

UBC
Economics 567

January 28, 2021
1 Introduction

2 Demand in product space

3 Demand in characteristic space
   Early work
   Model
   Model
   Estimation and identification
      Aggregate product data
      Estimation steps
      Pricing equation
      Micro data

4 References
References

• **Reviews:**
  - Ackerberg et al. (2007) section 1 (these slides use their notation)
  - Aguirregabiria (2019) chapter 2
  - Reiss and Wolak (2007) sections 1-7, especially 7

• **Classic papers:**
  - Berry (1994)
  - Berry, Levinsohn, and Pakes (1995)
Section 1

Introduction
• Typical market for consumer goods has many differentiated, but similar products, e.g.
  - Cars
  - Cereal
• Differentiated products are a source of market power
• Having many products can result in many parameters creating estimation difficulties and requiring departures from textbook demand and supply models
Motivation

- Counterfactuals that do not change production technology
  - Mergers
  - Tax changes
- Effects of new goods
- Cost-of-living indices
- Product differentiation and market power
  - Cross-price elasticities
Section 2

Demand in product space
Demand in product space

- $J$ products, each treated as separate good
- Classical demand,

$$q_1 = D_1(p_1, ..., p_J, z_1, \eta_1; \beta_1)$$

$$\vdots = \vdots$$

$$q_J = D_J(p_1, ..., p_J, z_J, \eta_J; \beta_J),$$

and supply (firms’ first-order conditions for prices):

$$p_1 = g_1(q_1, ..., q_J, w_1, \nu_1; \theta_1)$$

$$\vdots = \vdots$$

$$p_J^d = g_J(q_1, ..., q_J, w_J, \nu_J; \theta_J),$$

where
Demand and supply of differentiated products

Paul Schrimpf

Introduction

Demand in product space

Demand in characteristic space

Early work
Model
Model Estimation and identification
Aggregate product data
Estimation steps
Pricing equation
Micro data

References

References

Demand in product space 2

- $p_j = \text{price}$
- $q_j = \text{quantity}$
- $z_j = \text{observed demand shifter}$
- $\eta_j = \text{unobserved demand shock}$
- $\beta_j = \text{demand parameters}$
- $w_j = \text{observed supply shifter}$
- $\nu_j = \text{unobserved supply shock}$
- $\theta_j = \text{supply parameters}$

- $D_j$ typically parametrically specified, e.g.

$$\ln q_j = \beta_{j0} + \beta_{j1} p_1 + \cdots + \beta_{jj} p_J + \beta_{jy} \ln y + Z_1 \gamma + \nu_j$$
Demand in product space

- Use reduced form to find instruments

\[ q_1 = \Pi^q_1(Z, W, \nu, \eta; \beta, \theta) \]

\[ \vdots = \vdots \]

\[ q_J = \Pi^q_J(Z, W, \nu, \eta; \beta, \theta) \]

\[ p_1 = \Pi^p_1(Z, W, \nu, \eta; \beta, \theta) \]

\[ \vdots = \vdots \]

\[ p_J = \Pi^p_J(Z, W, \nu, \eta; \beta, \theta) \]

- Cost shifters of product \( j \) excluded from demand and supply of product \( k \), but in reduced form
  - Cost data often not available
  - If available, unlikely to be product specific
- Attributes of other products
  - Hausman (1996) uses prices of other products
  - Hard to justify, especially with prices
Demand in product space

- Advantages of product space:
  - Flexible substitution patterns
  - Does not require detailed product attribute data

- Problems with product space:
  1. Representative agent and aggregation issues
      - With heterogeneous preferences, aggregate market demand need not meet restrictions on individual demand derived from economic theory
      - Cannot use restrictions easily to improve estimates
      - Can use simulation to aggregate (Pakes, 1986)
  2. Too many parameters, $O(J^2)$
      - Can limit by restricting cross-price elasticities, e.g. Pinkse, Slade, and Brett (2002)
  3. Too many instruments needed, $J$
  4. Cannot analyze new goods
Section 3

Demand in characteristic space
Demand in characteristic space

• Motivation:
  • Why do firms differentiate products?
  • Because consumers have heterogeneous tastes for product characteristics
    • E.g. cars: tastes for size, safety, fuel efficiency, etc

• Main idea: model consumer preferences for characteristics and treat products as bundles of characteristics

• Early work: Lancaster (1971), McFadden (1973)

• Key extension to early work: Berry, Levinsohn, and Pakes (1995)
Early work in characteristic space

- Consumer chooses one or none of $J$ products
- Utility of consumer $i$ from product $j$
  \[ u_{ij} = x_j \beta + \epsilon_{ij} \]
  with $\epsilon_{ij}$ iid across $i$ and $j$ (usually Type I extreme value)
- Implies aggregate demand (for logit)
  \[ q_j = \frac{\exp(x_j \beta)}{1 + \sum_{k=1}^{J} \exp(x_k \beta)} \]
- Problem: restrictive substitution “independence of irrelevant alternatives”
  - Two goods with the same shares have the same cross price elasticities with any third good (think about a luxury and bargain good with equal shares)
  - Goods with same shares should have same markups
- Solution: add heterogeneity in $\beta$ and/or allow correlation across $j$ in $\epsilon_{ij}$
Model 1

- Consumers $i$, goods $j$, markets $t$
- Utility: (include good $0 =$ buy nothing)

\[ u_{ijt} = U(\tilde{x}_{jt}, \xi_{jt}, Z_{it}, v_{it}, p_{jt}; \theta) \]

\[ x_{jt} = (\tilde{x}_{jt}, p_{jt}) \in \mathbb{R}^K, z_{it} \in \mathbb{R}^R, v_{it} \in \mathbb{R}^L \]

- Choose $j$ if $u_{ijt} > u_{ikt} \forall k \neq j$
• Usually $U(\cdot)$ linear:

$$u_{ijt} = \begin{bmatrix} 1 & \mathbf{x}_{jt} & \mathbf{\theta}_{it} & 1 \end{bmatrix} + \mathbf{\xi}_{jt} + \mathbf{\epsilon}_{ijt}$$

$$= \bar{\theta} + \theta^o \mathbf{z}_{it} + \theta^u \mathbf{v}_{it}$$

for $j = 1 \ldots J$ and normalize $u_{i0t} = 0$

• Assume $\mathbf{\epsilon}_{ijt}$ i.i.d. double exponential

• Assume $\mathbf{v}_{it} \sim f_{\mathbf{v}}(\cdot; \theta)$, e.g. independent normal

• Write as product specific + observed interactions + unobserved interactions

$$u_{ijt} = \begin{bmatrix} \delta_j \end{bmatrix} + \mathbf{x}_{jt} \begin{bmatrix} \theta^o \end{bmatrix} \mathbf{z}_{it} + \mathbf{x}_{jt} \begin{bmatrix} \theta^u \end{bmatrix} \mathbf{v}_{it} + \mathbf{\epsilon}_{ijt}$$

$$= \mathbf{x}_{jt} \bar{\theta} + \mathbf{\xi}_{jt}$$
Endogeneity

- Usually assume $E[v_{it}|x_{jt}, z_{it}] = 0$ and $E[\epsilon_{ijt}|x_{jt}, z_{it}] = 0$
- Not interested in counterfactuals with respect to changes in $z_{it}$, so can treat as residual, i.e.

$$v_{it} = \theta_{it} - E[\theta_{it}|z_{it}]$$

- Market average $v_{it}$ or $\epsilon_{ijt}$ plausibly correlated with $p_{jt}$ or other product characteristics, but this correlation absorbed into $\xi_{jt}$ and/or market fixed effects
• Problem is $\xi_{jt}$
  • Prices and other flexible product characteristics must be correlated with $\xi_{jt}$
  • If $\xi_{jt}$ serially correlated, then likely also correlated with inflexible product characteristics
• Need instrument, $w_{jt}$ such that $E[\xi_{jt} | w_{jt}] = 0$
  • Cost shifters
  • Characteristics of other products
Estimation and identification

- Depends on data:
  - Aggregate product market shares and characteristics
  - Individual characteristics and choices
- Additional assumptions:
  - Use supply and equilibrium assumptions to get a pricing equation
Aggregate data 1

- Often only have data on product characteristics and market shares
- Maybe also distribution of some individual characteristics for each market (e.g. income and education from CPS or census)
- Instrument $w$ such that $E[\xi_j|w] = 0$
- Distribution of $\nu \sim f_\nu(\cdot; \theta_\nu)$
  - Combination of estimated market level distribution of observed individual characteristics and parametric distributions of unobserved individual characteristics
  - e.g. $\nu_{it} = (\text{educ}_{it}, \text{income}_{it}, e_{it})$

$$F_{\nu,t}(s, y, e; \theta_\nu) = \hat{F}_t(s, y) \Phi \left( \frac{e - \theta_\nu^\mu}{\theta_\nu^\sigma} \right)$$

$\hat{F}_t(s, y)$ estimated from CPS or other similar data set
• Assume $\varepsilon_{ijt} \sim$ double exponential (aka Gumbel or type I extreme value) as in logit
  • Computationally convenient, but other distributions feasible too
Estimation outline

- Estimate $\theta$ from moment condition
  \[ E[\xi(\cdot; \theta) | w] = 0 \]

- Where $\xi(\cdot; \theta)$ is such that model predicted market shares $=$ observed market shares$^1$
  1. Compute shares given $\theta$, $\sigma(\cdot; \theta, \delta)$
  2. Find $\delta(\cdot; \theta) = x_{jt} \bar{\theta} + \xi(\cdot; \theta)$ such that observed shares, $s_{jt}$ $=$ model shares, $\sigma(\cdot; \theta, \delta)$, then
    \[ \xi(\cdot; \theta) = \delta(\cdot; \theta) - x_{jt} \bar{\theta} \]

---

$^1$In this slide $\cdot$ means the data. I will leave the $\cdot$ out of the notation in subsequent slides. I will also leave out $t$ subscripts.
Computing model shares

- Integrate over $\nu$

$$\sigma_j(\theta, \delta) = \int \frac{\exp(\delta_j + x_j \theta^u \nu)}{1 + \sum_{k=1}^{j} \exp(\delta_k + x_k \theta^u \nu)} dF_\nu(\nu)$$

- Integral typically has no closed form, so compute numerically, usually by Monte Carlo integration

$$\sigma_j(\theta, \delta) = \sum_{r=1}^{N_s} \frac{\exp(\delta_j + x_j \theta^u \nu_r)}{1 + \sum_{k=1}^{j} \exp(\delta_k + x_k \theta^u \nu_r)}$$

where $\nu_r$ are $N_s$ random draws from $f_\nu$

- Issues about how best to compute integral — simulation vs quadrature, type of simulation (Skrainka and Judd, 2011)
- Simulation (more generally approximation) of integral affects distribution of estimator
Solving for $\delta$ and $\xi$

- Want $\delta$ s.t. $\sigma_j(\theta, \delta) = \hat{s}_j$

- Berry, Levinsohn, and Pakes (1995) show

\[
T(\delta) = \delta + \log(\hat{s}_j) - \log(\sigma_j(\theta, \delta))
\]

is a contraction

- Unique fixed point $\delta$ such that
  \[\delta = \delta + \log(\hat{s}_j) - \log(\sigma_j(\theta, \delta)), \text{ i.e. } \hat{s}_j = \sigma_j(\theta, \delta)\]

- Can compute $\delta(\theta)$ by repeatedly applying contraction (in theory and practice often faster to use other method)

- $\xi_j(\theta) = \delta_j(\theta) - x_j \bar{\theta}$

- Important identifying assumption: only $\theta$ s.t. $\bar{\xi}_j(\theta) = \bar{\xi}_j^0$ is true $\theta_0$
Estimating $\theta$

- Conditional moment restriction $E[\xi_j(\theta)|w] = 0$
- Empirical unconditional moments:

$$G_{J,T,N,N_s} = \frac{1}{JT} \sum_{j=1}^{J} \sum_{t=1}^{T} \xi_{jt}(\theta)f(w_t)$$

where

- $f(w) = \text{vector of function of } w$
- $J = \text{number of products}$
- $T = \text{number of markets}$
- $N = \text{number of observations in each market underlying } \hat{s}_j$
- $N_s = \text{number of simulations}$

- Asymptotic properties (consistency, distribution), depend on which of $J, T, N,$ and $N_s$ are $\rightarrow \infty$, see Berry, Linton, and Pakes (2004)
- Reynaert and Verboven (2014): using optimal instruments greatly improves efficiency and stability
Pricing equation 1

- More moments give more precise estimates
- Assumption about form of equilibrium allows use of firm first order condition (pricing equation) as additional moment
- Nash equilibrium in prices
- Log linear marginal cost

\[ \log mc_j = r_j \theta^k + \omega_j \]

- \( r_j \) = observed product characteristics, input prices, maybe quantity, etc
- \( \omega_j \) = unobserved productivity, possibly endogenous
- Firm \( f \) producing set of product \( J_f \),

\[ \max_{p_j: j \in J_f} \sum_{j \in J_f} (p_j - C_j(\cdot)) M_s_j(\cdot, p) \]
Pricing equation 2

- First order condition:

\[ \sigma_j(\cdot) + \sum_{l \in J_f} (p_l - mc_l) \frac{\partial \sigma_l(\cdot)}{\partial p_j} = 0 \]

- Collect as

\[ s + (p - mc)\Delta = 0 \]

- Rearrange and use log linear marginal cost

\[ \log(p - \Delta^{-1}\sigma) - r\theta^c = \omega(\theta) \]

- Conditional moment restriction \( E[\omega(\theta)|w] = 0 \)

- Add empirical moments to \( G, \frac{1}{JT} \sum_{jt} \omega_{jt}(\theta)f(w_t) \)
Micro data

- **Berry, Levinsohn, and Pakes (2004)**
- **Data on individual choices and characteristics**

\[ u_{ijt} = \delta_j + x_{jt} \theta^o Z_{it} + x_{jt} \theta^u v_{it} + \epsilon_{ijt} \]

\[ = x_{jt} \bar{\theta} + \xi_{jt} \]

- Random coefficients discrete choice model, so can identify and estimate \( \delta, \theta^o, \) and \( \theta^u \) without assumptions about \( \xi \) and \( x \)
  - Ichimura and Thompson (1998) give conditions for nonparametric identification of random coefficients binary choice models
  - Estimate by MLE or (usually) GMM
  - Still need \( \bar{\theta} \) for price elasticities, etc

\[ \delta_j = x_{jt} \bar{\theta} + \xi_{jt} \]

- Use IV
- Use IV with a pricing equation
Section 4

References


