Demand estimation part I+part II

Why estimating demand?

Marketing service:

- For the pricing strategy.
  Usually we can only charge one price for the product: need to fix it right! If the price too low, we "give it away" to many consumers at a price below their reservation price. If price is too high, we lose a lot of consumers.
- For new goods it is important to predict to what extent consumers are willing to substitute their previous favorite brands by the new brand.
- Study of consumers' taste formation: habits, brand loyalty, effect of advertising.

Why estimating demand?

For policy evaluation:

- If we study the effects of a tax change, we want to distinguish the effects on consumer surplus and on firms' profits.
- Also, the tax on a specific product can drive some consumers to substitute it by other products.
- We are interested in measuring this substitution effects via the estimation of a demand system.

Why estimating demand?

Competition authority:

- To assess market power of firms
- Relationship between price sensitivity and market power, the monopoly case:

$$\frac{p - mc}{p} = \frac{-1}{\eta}$$

. In Cournot competition: Individual mark-up

$$\frac{p - mc_i}{p} = \frac{-s_i}{\eta}$$

Average mark-up:

$$\sum_i s_i \frac{p - mc_i}{p} = \sum_i \frac{-s_i^2}{\eta} = -\frac{HHI}{\eta}$$

Why estimating demand?

Competition authority:

- estimation of demand is also used to define market: "Are two products close enough so they belong to the same market?"
- Crucial for horizontal merger analysis: measure current market power and predict post-merger market power and price increase.
- For merger analysis: the profitability of a horizontal merger depends crucially on the degree of substitution between the products that these firms produce

What is demand curve?

- A demand curve or demand function indicates the quantity of a good that can be sold, given various elements of the marketing mix or some other environmental or "macro" variables:
    - Price
    - Advertising
    - Availability of coupons
    - Income
    - Housing value
    - Season (Christmas, Easter, ...)
- Both own and competitor's actions will generally influence the demand curve.

Interpretation of demand curve

Demand curve as derived from the utility function of the consumer:

- Underlying assumption: consumer can buy any amount of the product (0.6 or 7.1 units)
- $\Rightarrow$ Continuous quantity demanded
- In most of purchase situations, the consumer purchases in discrete amounts, i.e. one, two, three, etc. units of the product
- $\Rightarrow$ Discrete quantity demanded.
- When many consumers (aggregate demand): discrete quantities become continuous

Often, the choice is between one unit or no purchase at all (e.g cars)

Then, we will use probabilities of purchase

Interpretation of demand curve

In terms of reservation prices :

- Reservation price is the maximum price a consumer is willing to pay for a good
- If the price offered is below or equal to the reservation price, the consumer will buy the product
- Different consumers have different willingness to pay or reservation price for the product
- Aggregate or market demand comes from summing up the demand of each consumer
- At some price P*, a certain fraction of the consumers in the market will buy
- The number of consumers who buy is the quantity of the product demanded Q(P*)
- What determines the level of demand and price sensitivity is how the reservation prices vary across consumers

Price Elasticities

- A useful measure of the response of sales to price is the price elasticity. The price elasticity of demand is defined as the percentage change in the quantity demanded given a percentage change in price:

$$\eta = \frac{\Delta Q / Q}{\Delta P / P}$$

- The price elasticity is typically negative (sometimes, the absolute value of the price elasticity is reported)
- Price elasticity useful to understand the nature of goods:
  - $|\eta| > 1$: demand elastic
  - $|\eta| < 1$: demand inelastic
  - $\eta > 0$: Giffen good

Price Elasticities

What drives the size of the elasticity? A lot of things!

- Availability of substitutes
- Extent of substitutability
- Awareness of prices (search costs?)
- Substitution also occurs over time (storage, purchase acceleration) and over different retail outlets/points of distribution
- Extent of substitutability depends on quality perceptions!
- Possibilities for substitution may vary over time due to technological advances

Demand for homogenous products

Example of homogenous products:

- Raw materials, commodities such as oil, sugar
- Finished products such as paper

- Typical industry data consist of quantities and prices:
  Data $=\{Q_t, P_t\}_{t=1\ldots T}$

- Data are generated by demand and supply functions

- How to infer the demand function and supply function from the observation of equilibrium prices and quantities ?

- Approach is more reduced-form than structural (we look for demand shifters instead of deriving demand function from consumers utility)

The linear Demand Model

We could start out with a simple linear model

$$Q = a + bP + \epsilon$$

- Q = sales in units (not revenue)
- P = price (in €)
- b is a parameter, which we would expect to be negative
- The error term represents demand shocks but also the effects of omitted factors on demand
  Problems with this linear model:
- Does not fit data well
- Effect of price change (ΔP) is the same regardless of the level of the price

The Log-linear Demand Model

A better model, and the basis of applied demand analysis is the log-linear demand model:

$$Q = AP^{-\eta}$$

- A and η are the "demand parameters"
- A: "intercept": drives base sales;
  η: "Slope": drives curvature response to price changes
- Each specific demand function will have its own parameters, such as $A = 500$ and $\eta = 2.2$
- Note: the effect of a change in price depends on the level of price
- At a high price level, a given ΔP will have a smaller effect on sales than at a lower price level.
- This model is also called multiplicative model

The Log-linear Demand Model

$$Q = AP^{-\eta}$$

Now take logs on both sides:

$$\log(Q) = \log(A) - \eta \log(P) = a - \eta p$$

Why is that useful?

- Parameters a and η enter linearly, and can therefore be estimated using standard regression analysis
- The parameter η has a simple and straightforward interpretation – it is the absolute value of the price elasticity of demand!

Estimation of the model

Introduce a multiplicative error term: $Q = AP^{-\eta}e^{\epsilon}$

- The error term $exp(\epsilon)$ is distributed only over positive values
- To estimate this model, we just need to take logs:

$$\log(Q_t) = \alpha + \beta \log(P_t) + \epsilon_t$$
$$\alpha = \log(A)$$
$$\beta = -\eta$$

- Estimation via OLS if price is exogenous: $\mathbb{E}(\epsilon | log(P)) = 0$

Note: We can add X_t, variables that may affect demand such as

- Marketing expenses, income, seasonal dummies...

Endogeneity

Main issue: quantity and price simultaneously determined

Price is endogenous

Instruments:

- Cost shifters
- Come from the specification of the supply function

Supply Side

Specification of marginal cost function:

$$mc_t = w_t\gamma + \lambda Q_t + u_t$$

Where:

- mc_t marginal cost in market t
- w_t cost shifters (input prices, technology shifters)
- Q_t quantity produced
- u_t unobserved (to the econometrician) cost shifters

- If perfect competition: $P_t = mc_t$

- We estimate:

$$\begin{cases} \log(Q_t) = X_t\alpha + \beta\log(P_t) + \epsilon_t \\ P_t = w_t\gamma + \lambda Q_t + u_t \end{cases}$$

We estimate:

$$\begin{cases} \log(Q_t) = X_t\alpha + \beta\log(P_t) + \epsilon_t \\ P_t = w_t\gamma + \lambda Q_t + u_t \end{cases}$$

Demand shifters: x_n and ϵ_n (e.g: change in income, availability of complements goods)

Cost shifters, w_n and u_n (e.g: wages, capacity expansion, firm entry)

Shifters are exogenous, and play the role of instruments

Endogenous variables are P and Q

Background

Most straightforward approach is to specify a demand function directly:

$$q = D(p, r, \varepsilon)$$

- q is a $J \times 1$ vector of quantities
- p is a $J \times 1$ vector of prices
- r is a vector of exogenous variables
- ε is a $J \times 1$ vector of random shocks.

Early work focused on how to specify $D(\cdot)$ in a way that was both flexible and consistent with economic theory.

Linear Expenditure model (Stone, 1954), Rotterdam model (Theil, 1965; Barten 1966), Translog model (Christensen, Jorgenson, and Lau, 1975), Almost Ideal Demand System (Deaton and Muellbauer, 1980), ...

Issues

Issues for many cases considered in IO:

"The too many parameters problem". As the number of options, J, becomes large there are too many parameters to estimate

- Even if $D(p, r, \varepsilon) = Ap + \varepsilon$, where A is $J \times J$ matrix of parameters, implies J^2 parameters to be estimated.
- With a more flexible functional form, the problem is even greater.

Heterogeneity. for some applications we would like to explicitly model and estimate the distribution of heterogeneity.

- Above approach, generally, does not let us do this
- Presence of heterogeneity doesn't invalidate aggregate demand but
  - should care in imposing the restrictions of economic theory
  - for example if preferences are of Gorman form (Gorman, 1959)

Issues

Explicit modeling of specific consumer behavior. easier when we start with an explicit model of consumer behavior and aggregate to market level demand.

New Goods. This demand system does not easily allow us to predict the demand for new goods.

Estimation. need to include highly colinear prices, and also instrument for them

Approaches

- Modeling of demand either in *product space* or in *characteristics space*
- Implicit vs. explicit modeling of heterogeneity
- Modeling choices and trade-offs to be addressed:
  - aggregation across products
  - symmetry across products
  - weak separability across products and multi stage budgeting
  - models in characteristics space and discrete choice

Modeling Approach in IO

- Random utility model underlying decisions of demand
- Dimensionality reduction with characteristics approach
- Consumer choice will depend on the best alternative among a given choice set

Preliminaries to the logit model

The set of alternatives, called the choice set must fulfill 3 conditions:

- The alternatives must be mutually exclusive from the decision maker's perspective.
- The choice set must be exhaustive.
- The number of alternatives is finite.

Models in Characteristics Space

- Previous examples are models in product space
- Models in characteristics space: the utility function is a function of the attributes of the product (Gorman, 1956/1980, Lancaster, 1966)
  - Connections to hedonics
  - Usually implemented as discrete choice, but does not have to be
- If the product identity is a characteristic, then approach nests the product space approach
- Indirect utility model at the level of the decision maker
- Modeling in products spaces using characteristics space for flexibility (Dubois, Griffith, Nevo, 2014)

Models in Characteristics Space

- Indirect utility function of consumer i, from product j in market t:

$$U(x_{jt}, \xi_{jt}, I_i - p_{jt}, \tau_i; \theta)$$

where

x_{jt} – $1 \times K$ vector of observed product characteristics
ξ_{jt} – unobserved (by us) product characteristic
τ_i – individual characteristics
I_i – income

- ξ_{jt} will play an important role
  - realistic (inability of observed characteristics to capture the essence of the product)
  - will act as a "residual" (why predicted shares do not fit exactly)
  - potentially implies endogeneity

Choice and Normalizations

- Consumer chooses j if

$$U(x_{jt}, \xi_{jt}, I_i - p_{jt}, \tau_i; \theta) \geq U(x_{kt}, \xi_{kt}, I_i - p_{kt}, \tau_i; \theta)$$

  for all $k \neq j$.
- For what follows utility is ordinal and not cardinal, and therefore is invariant to affine transformation
- This means that we typically have 2 normalizations
- In what follows (will make more sense below)
  - normalize the variance of one component to one
  - normalize the utility from the outside good to zero

Logit Model

- A common model of this class is the Logit model

Individual random utility for the $J+1$ goods can be written ($j=0,..J$)

$$u_{ij} = x_j\beta - \alpha p_j + \xi_j + \varepsilon_{ij}$$

where ε_{ij} is a stochastic term
(equivalent to $u_{ij} = x_j\beta - \alpha(y_i - p_j) + \xi_j + \varepsilon_{ij}$ because all variables non specific to alternative cancel out)

- 2 views of ε_{ij}:
  - utility is deterministic, but the choice process itself is probabilistic (Tversky, 1972)
  - utility is deterministic, but ε_{ij} captures the researcher's inability to formulate individual behavior precisely
- In a way all what ξ_j is doing is changing the mean of ε_{ij}, by j.

Logit Model

- Assume individual chooses the product that maximizes random utility and derive individual choice probabilities
- Aggregate market shares can be derived from individual choice probabilities
- Here, all consumers have same valuations for the characteristics and price, we can write

  $$u_{ij} = \delta_j + \varepsilon_{ij}$$

  where

  $$\delta_j = x_j \beta - \alpha p_j + \xi_j$$

  is the mean or common part of consumers' utility.
- The mean utility of the outside good is normalized to 0 so $\delta_0 = 0$.
- In the logit model, consumers only differ in ε_{ij}.

Logit Model

- ε_{ij} is modeled as an i.i.d. random variable with an extreme value distribution

$$F(\varepsilon_{ij}) = \exp(-\exp(-\varepsilon_{ij}))$$

- Choice probability

$$\begin{aligned} P_{ij} &= P\left(u_{ij} \geq u_{ik}, \forall k \neq j\right) \\ &= P\left(\varepsilon_{ik} - \varepsilon_{ij} \leq \delta_j - \delta_k, \forall k \neq j\right) \end{aligned}$$

- The difference between two extreme value variables is distributed logistic such that

$$F(\varepsilon_{ij} - \varepsilon_{ik}) = \frac{\exp\left(\varepsilon_{ij} - \varepsilon_{ik}\right)}{1 + \exp\left(\varepsilon_{ij} - \varepsilon_{ik}\right)}$$

Logit Model

- Logit choice probabilities

$$
\begin{aligned}
P_{ij} &= P\left(u_{ij} \geq u_{ik}, \forall k \neq j\right) = P\left(\varepsilon_{ik} - \varepsilon_{ij} \leq \delta_j - \delta_k, \forall k \neq j\right) \\
&= P\left(\varepsilon_{ik} \leq \varepsilon_{ij} + \delta_j - \delta_k, \forall k \neq j\right) \\
&= \int P\left(\varepsilon_{ik} \leq \varepsilon_{ij} + \delta_j - \delta_k, \forall k \neq j | \varepsilon_{ij}\right) dF\left(\varepsilon_{ij}\right) \\
&= \int \prod_{k \neq j} \exp(-\exp(-\varepsilon_{ij} - \delta_j + \delta_k)) dF\left(\varepsilon_{ij}\right) \\
&= \int \prod_{k \neq j} \exp(\text{-}\exp(\text{-}\varepsilon_{ij}\text{-}\delta_j + \delta_k)) \exp(\text{-}\exp(\text{-}\varepsilon_{ij})) \exp(\text{-}\varepsilon_{ij}) d\varepsilon_{ij}
\end{aligned}
$$

- Details next page

Logit Model

$$
\begin{aligned}
P_{ij} &= \int_{-\infty}^{+\infty} \left(\prod_{k \neq j} \exp(-\exp(-\varepsilon_{ij} - \delta_j + \delta_k)) \right) \exp(-\exp(-\varepsilon_{ij})) \exp(-\varepsilon_{ij}) d\varepsilon_{ij} \\
&= \int_{-\infty}^{+\infty} \exp(-\sum_k \exp(-\varepsilon_{ij} - \delta_j + \delta_k)) \exp(-\varepsilon_{ij}) d\varepsilon_{ij} \\
&= \int_{-\infty}^{+\infty} \exp(-\exp(-\varepsilon_{ij}) \sum_k \exp(-\delta_j + \delta_k)) \exp(-\varepsilon_{ij}) d\varepsilon_{ij} \\
&= \int_{0}^{+\infty} \exp(-\theta \sum_k \exp(-\delta_j + \delta_k)) d\theta \text{ (change of variable)} \\
&= \left[\frac{\exp(-\theta \sum_k \exp(-\delta_j + \delta_k))}{-\sum_k \exp(-\delta_j + \delta_k)} \right]_0^{+\infty} = \frac{1}{\sum_{k=0}^{J} \exp(\delta_k - \delta_j)} \\
P_{ij} &= \frac{\exp \delta_j}{1 + \sum_{k=1}^{J} \exp \delta_k} \in (0, 1)
\end{aligned}
$$

Logit Model

- IIA property

$$\frac{P_{ij}}{P_{ij'}} = \frac{\frac{\exp \delta_j}{1+\sum_{k=1}^{J} \exp \delta_k}}{\frac{\exp \delta_{j'}}{1+\sum_{k=1}^{J} \exp \delta_k}} = \exp\left(\delta_j - \delta_{j'}\right)$$

- What does it mean?

Independence of the Irrelevant Alternatives (IIA)

Consider the well-known red-bus-blue-bus problem. A person has the choice of going to work by car or taking a blue bus. We assume that the representantive utilities of the two modes are the same, and therefore $P_c = P_{bb} = 1/2$. Therefore the ratio of probabilities is one: $P_c/P_{bb} = 1$ The red bus is now introduced and the traveler consider the red bus to be the same like the blue bus.$P_{rb}/P_{bb} = 1$. But in the logit model the model P_c/P_{bb} is the same whether or not a new bus is introduced. The only probabilities for which $P_c/P_{bb} = 1$ and $P_{rb}/P_{bb} = 1$ are $P_c = P_{bb} = P_{rb} = 1/3$, which are the probabilities that the logit model predicts.

Independence of the Irrelevant Alternatives (IIA)

What would be more realistic to expect?

- the original probability of taking the bus to be split between the two buses.
- $P_c = 1/2$ and $P_{bb} = P_{rb} = 1/4$.
- the ratio P_c/P_{bb} should actually change and not remain constant as required by the logit model.

Logit Model

- Consumer surplus

$$CS = \frac{1}{\alpha} \max_j u_{ij}$$

where α is the marginal utility of income

- Then

$$\begin{aligned} E\left[CS\right] &= \frac{1}{\alpha} E\left[\max_j u_{ij}\right] \\ &= \frac{1}{\alpha} E\left[\max_j \left(\delta_j + \varepsilon_{ij}\right)\right] \\ &= \frac{1}{\alpha} \ln \sum_{k=1}^{J} \exp \delta_k + C \end{aligned}$$

- Change in consumer surplus between 0 and 1

$$\Delta E\left[CS\right] = \frac{1}{\alpha} \left(\ln \sum_{k=1}^{J^1} \exp \delta_k^1 - \ln \sum_{k=1}^{J^0} \exp \delta_k^0 \right)$$

Logit Model

- Choice probabilities

$$P_{ij}\left(\alpha, \beta, \xi\right) = \frac{\exp\left(x_j\beta - \alpha p_j + \xi_j\right)}{1 + \sum_{k=1}^{J} \exp\left(x_k\beta - \alpha p_k + \xi_k\right)}$$

  and for the outside good

$$P_{i0}\left(\alpha, \beta, \xi\right) = \frac{1}{1 + \sum_{k=1}^{J} \exp\left(x_k\beta - \alpha p_k + \xi_k\right)}$$

- McFadden (1974) proved that the log-likelihood function with these choice probabilities is globally concave in parameters (α, β, ξ).

Logit Model Estimation

- Estimation by Maximum Likelihood with N observations (decision makers) is possible if sample is exogenously drawn

$$\max_{\alpha,\beta,\xi} \sum_{i=1}^{N} \sum_{j} d_{ij} \ln P_{ij}\left(\alpha, \beta, \xi\right)$$

where $d_{ij} = 1$ if i chooses j and 0 otherwise

Logit Model Estimation: Aggregate Data

- The aggregate market share function is simply equal to the individual choice probability, as a function of the vector δ.
- With M consumers, the observed market share (relative to the potential market) is simply

$$s_j = \frac{q_j}{M}$$

- Equating market share functions to observed market shares (in vector form)

$$P(\alpha, \beta, \xi) = s$$

  but error terms ξ_j enter non linearly through the mean utilities δ.

Logit Model Estimation: Aggregate Data

- Berry (1994) showed how to invert the demand system easily to solve for the mean utility vector
- We have

$$s_j = \frac{\exp\left(x_j\beta - \alpha p_j + \xi_j\right)}{1 + \sum_{k=1}^{J} \exp\left(x_k\beta - \alpha p_k + \xi_k\right)}$$

$$s_0 = \frac{1}{1 + \sum_{k=1}^{J} \exp\left(x_k\beta - \alpha p_k + \xi_k\right)}$$

- So that

$$\begin{aligned} \frac{s_j}{s_0} &= \exp\left(x_j\beta - \alpha p_j + \xi_j\right) \\ \ln\left(\frac{s_j}{s_0}\right) &= x_j\beta - \alpha p_j + \xi_j \end{aligned}$$

- We obtain a standard linear regression model

Logit Model Estimation: Aggregate Data

- We observe quantities demanded and not market shares, but if we know the market size M then

$$\ln\left(\frac{q_j\,/\,M}{q_0\,/\,M}\right) = \ln\left(\frac{q_j}{M - \sum_{k=1}^{J} q_k}\right)$$

- Estimation with different markets t (or periods):

$$\ln\left(\frac{s_{jt}}{s_{0t}}\right) = x_{jt}\beta - \alpha p_{jt} + \xi_{jt}$$

- Time (market) dummies allow to capture unknown market size.

Logit Model Estimation: Aggregate Data

- Linear equation

$$\ln\left(\frac{s_{jt}}{s_{0t}}\right) = x_{jt}\beta - \alpha p_{jt} + \xi_{jt}$$

  The price may be correlated with error term, such that OLS would lead to an estimate of α which is biased towards zero.

- Solution: 2SLS
- Instruments to identify the parameters:
  - cost side variables (not always available with product level data)
  - characteristics of competitors (see BLP 1995): x_j and $\sum_{k \neq j} x_k$ and J and similar sums by firm
  - lagged variables or variables from other markets (if panel data)

Logit Model Estimation: Aggregate Data

- As
$$\frac{\partial s_j}{\partial p_j} = -\alpha s_j\left(1 - s_j\right) \text{ and } \frac{\partial s_k}{\partial p_j} = \alpha s_j s_k$$
- Elasticities have simple expressions
$$\eta_{jk} = \frac{\partial s_j}{\partial p_k}\frac{p_k}{s_j} = \begin{cases} -\alpha p_j(1 - s_j) & \text{if } j = k \\ \alpha p_k s_k & \text{otherwise} \end{cases}$$
- Problems:
    - Own price elasticities: market shares are small, so $\alpha(1 - s_j)$ is nearly constant and therefore the own-price elasticities are proportional to price. This is driven mostly by lack of heterogeneity
    - Cross-price elasticities: cross price elasticity w.r.t. a change in the price of product k is that same for all products such that $j \neq k$. This is driven by lack of heterogeneity and i.i.d.

Demand for differentiated products - characteristics approach

- Data: Typically market level (country, city,...) quantities and prices at brand level (from retailers, ACNielsen). Also need data on product characteristics. Data on consumer choices or characteristics are nice but not neccesary. Data from several geographical markets also nice, but not neccessary. No data on mc are typically used (some input prices are nice to have but, again, are not neccessary.
- Demand: Utility of consumers is assumed to depend on product characteristics and individual taste parameters. Typically a discrete choice approach is used (McFadden). See Pinkse-Slade-Brett (Econometrica, 2002) for an alternative approach.
- Supply: Typical assumption is one shot Bertrand. Supply and demand sometimes estimated jointly, sometimes demand estimated and prices and markups simply backed out.

Demand for differentiated products - Berry 1994

Divide products into G mutually exclusive nests (i.e. sparkling and still water) with the outside good as one nest.

$$u_{ij} = \delta_j + \zeta_{ig} + (1-\sigma)\,\epsilon_{ij}$$

where ϵ_{ij} and $\zeta_{ig} + (1-\sigma)\,\epsilon_{ij}$ are again distributed iid with a type I extreme value distribution. ζ_{ig} measures the common component in valuation of products withing each nest. $0 \leq \sigma < 1$ measures the correlation of tastes within nests.

$$\ln(s_j) - \ln(s_o) = \underbrace{x_j\beta - \alpha p_j + \sigma \ln(s_{j|g}) + \xi_j}_{\equiv \delta_j}$$

where $s_{j|g}$ is the market share of product j in nest g.

Demand for differentiated products -neste logit, Berry 1994

Substitution patterns are now richer and

$$\frac{\partial s_j}{\partial p_k}\frac{p_k}{s_j} = \begin{cases} \alpha p_j(s_j - 1/(1-\sigma) + \sigma/(1-\sigma)s_{j|g}) \text{ if } k = j \\ \alpha p_k(s_k + \sigma/(1-\sigma)s_{k|g}) \text{ if } k \neq j \text{ and same nest} \\ \alpha p_k s_k \text{ if } k \neq j \text{ and different nests} \end{cases}$$

Nested logit frequently used - if nesting is appropriate perceived to generate plausible substitution patterns.

Demand for differentiated products -neste logit, Random coefficient models

Random coefficient models allow individuals to vary in their valuation of characteristics

$$u_{ij} = x_{ij}\beta - \alpha_i p_j + \xi_j + \epsilon_{ij}$$

No closed form solutions. Analogous to mixlogit in STATA. Estimate demand by GMM. Key reference Berry, Levinsohn and Pakes (Econometrica 1995).

Endogeneity of price

Two sources of price endogeneity have been the focus of considerable attention. One is "cross-sectional" endogeneity of price across brands. An example may clarify – a Ferrari may have high product quality that is observed by both consumers and producers but for which we don't have an observable product characteristic in our data set (apart from potential dummies). The other worry is the time series endogeneity - positive demand shocks should raise price (even though the evidence on this is weaker than all referees think...). Common solutions:

- "Hausman instruments" Use prices in other local markets - assumption that they will feature same cost shocks but that demand shocks are local.

Endogeneity of price

- Product fixed effects. Used to control for cross-sectional endogeneity. Consider data from regional markets. Utility for consumer i of product j in market k at time t can be expressed as

$$u_{ijkt} = \underbrace{x_j\beta - \alpha p_{jkt} + \xi_j + \Delta\xi_{jkt}}_{\equiv \delta_{jkt}} + \epsilon_{ijkt}$$

where $\Delta\xi_{jkt}$ is the local deviation from mean valuation. In the logit case for instance we can then estimate the following:

$$\ln(s_{jkt}) - \ln(s_{okt}) = \underbrace{x_j\beta + \xi_j}_{\text{product f.e.}} - \alpha p_{jkt} + \Delta\xi_{jkt}$$

The correlation between price and the mean unobserved product quality that is the same across markets is now picked up by the product fixed effect.

Endogeneity of price

- "BLP instruments" The price that you set will depend on the characteristics of competing products. But the competing products have no direct effect on ξ_j under the assumption that characteristics are exogenous. Following this logic BLP 1995 suggested using functions of the characteristics of competing products as instruments.
- Cost shocks. Usually not enough variation in costs across products to use them as instruments. But in some settings exchange rates may work. Diesel price interacted with distance used in Friberg and Ganslandt (2007).

Supply Side

Profit Maximization Problem:

$$\max_{\{p_i\}_{i\in f}} \left\{ \sum_{i\in f} \left(p_i q_i(p) - c_i(q_i(p))\right) \right\}$$

First-Order Conditions (FOC)

$$\forall i \in f : q_i(p) + \sum_{j\in f} \left(p_j \frac{\partial q_j(p)}{\partial p_i} \right) - \sum_{j\in f} \left(mc_j(q_j) \frac{\partial q_j(p)}{\partial p_i} \right) = 0$$

Matrix Form of FOCs:

$$q + (\Omega \cdot D_p q)(p - mc) = 0$$

Here, q is the vector of quantities sold by each firm, Ω is the ownership matrix, $D_p q$ is the Jacobian matrix of partial derivatives of quantities with respect to prices, p is the vector of prices set by each firm, and mc is the vector of marginal costs for each firm.

Supply Side

The ownership matrix Ω accounts for which products are owned by the same firm. The "ones" in the matrix represent products that are owned by the same firm, and the "zeros" represent products owned by different firms.

- (One): If the firm owns both product j and product k, then Ω_{jk} is 1. This means that the firm's decision on the price of product j would affect the quantity sold of product k because they are under the same ownership.
- (Zero): If the firm does not own both products (i.e., product j and product k are owned by different firms), then Ω_{jk} is 0. There is no direct ownership link affecting how the price of product j would influence the quantity sold of product k.

Supply Side

So, in a simple case with products A, B, and C where A and B are owned by Firm 1, and C is owned by Firm 2, the matrix would look like this:

$$\Omega = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

- The first row says product A is owned by the same firm as product A (itself), product B (hence the 1 in the second column), but not product C (hence the 0 in the third column).
- The second row says product B is owned by the same firm as product A, product B (itself), but not product C.
- The third row says product C is not owned by the same firm as either product A or B, but it is owned by the same firm as product C (itself).

Supply Side

If there were more firms and products, the matrix would have more blocks of 1s along the diagonal, with each block representing a different firm and the products it owns. The size of each block corresponds to the number of products owned by that firm, and the off-block elements would be 0, indicating no ownership relationship between those products.

Using this ownership matrix, if we wanted to perform the Hadamard product with a matrix B of the same dimensions, the operation would be conducted element-wise. Let's suppose matrix B is a matrix of derivatives :

$$B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}$$

The Hadamard product $\Omega \circ B$ would be:

$$\Omega \circ B = \begin{bmatrix} 1 \cdot b_{11} & 1 \cdot b_{12} & 0 \cdot b_{13} \\ 1 \cdot b_{21} & 1 \cdot b_{22} & 0 \cdot b_{23} \\ 0 \cdot b_{31} & 0 \cdot b_{32} & 1 \cdot b_{33} \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & 0 \\ b_{21} & b_{22} & 0 \\ 0 & 0 & b_{33} \end{bmatrix}$$

Supply Side with conduct parameter

- There are J products indexed by j.
- The demand for product j is $q_j(p)$, where p is a $J \times 1$ price vector.
- The marginal cost for each product is constant and denoted as c_j.
- Each firm f owns a subset of products F_f.
- Each firm chooses the prices of its own products to maximize profits, defined as:

$$\pi_f(p) = \sum_{j \in F_f} (p_j - c_j) q_j(p) + \phi \sum_{j \notin F_f} (p_j - c_j) q_j(p)$$

If $\phi = 0$, each firm behaves non-cooperatively in the sense that they do not coordinate with other firms, but they still consider the impact of their pricing decisions across their own range of products.

In contrast, if ϕ were set to 1, this would imply that the firm is coordinating pricing not just internally but potentially across the entire market, akin to a cartel.

Supply Side with conduct parameter

- Firm 1 owns products A and B.
- Firm 2 owns product C.
- We'll use the conduct parameter ϕ to reflect the degree of competitive behavior or cooperation between products from different firms.
- The product-ownership matrix Θ, reflecting both the ownership and the competitive interaction, would be:

$$\Theta = \begin{bmatrix} 1 & 1 & \phi \\ 1 & 1 & \phi \\ \phi & \phi & 1 \end{bmatrix}$$

Premerger scenario

Initially, we have three firms with the following ownership of products:

- Firm 1 owns Product A and Product B.
- Firm 2 owns Product C.
- Firm 3 owns Product D.

The conduct parameter ϕ is set at 0.2, indicating a mild level of cooperation between the firms.
The pre-merger product-ownership matrix Θ^{pre} is defined as:

$$\Theta^{pre} = \begin{pmatrix} 1 & 1 & 0.2 & 0.2 \\ 1 & 1 & 0.2 & 0.2 \\ 0.2 & 0.2 & 1 & 0.2 \\ 0.2 & 0.2 & 0.2 & 1 \end{pmatrix}$$

Premerger scenario

The price is expressed as a function that includes marginal cost and a markup. The markup is inversely related to the price elasticities and the product-ownership matrix:

$$p = c - \{\Theta \circ \Delta(p)\}^{-1} q(p)$$

The equation can be used to uncover the premerger marginal cost vector based on premerger prices and estimated price elasticities of demand:

$$c^{pre} = p^{pre} + \{\phi^{pre} \circ \Delta(p^{pre})\}^{-1} q(p^{pre})$$

Postmerger scenario

After the merger of Firm 1 and Firm 2, the product ownership is realigned as follows:

- Merged Firm 1-2 owns Product A, Product B, and Product C.
- Firm 3 continues to own Product D.

The conduct parameter ϕ increases to 0.6 to reflect a higher degree of cooperation. The post-merger product-ownership matrix Θ^{post} becomes:

$$\Theta^{post} = \begin{pmatrix} 1 & 1 & 1 & 0.6 \\ 1 & 1 & 1 & 0.6 \\ 1 & 1 & 1 & 0.6 \\ 0.6 & 0.6 & 0.6 & 1 \end{pmatrix}$$

After a merger, both the product-ownership matrix (Θ) and the marginal cost vector (c) may change from premerger (Θ^{pre}, c^{pre}) to postmerger (Θ^{post}, c^{post}). Fixed-point iteration or other algorithms, such as the Newton method, can be applied to find the new equilibrium prices.