ECON 8000/9000 Empirical Energy Econ

Topic 14: Demand Estimation: Simple Multinomial Logit

Christy Zhou

April 22, 2026

Outline

- Introduction to Discrete Choice
- Introduction to Simple Multinomial Discrete Choice with Aggregate Data
- Output from Demand Estimation: Elasticities
- IV
- Key Practitioner Guide & Online Resource

Discrete Choice of Demand

Typical Scenarios

Often the goal for demand estimation is to:

- Undercover demand elasticity
- Learn about consumer preference for exogenous attributes

With key demand parameters uncovered, we can then simulate counterfactuals:

- **Shock 1: Price shock**
  What if some products receive a promo/tax/subsidy?
  How would consumers substitute towards or away from this product?
  What is the shift between subsidized products vs cannibalized sales within the same brand?
- **Shock 2: Quality shock** (or any attribute shock)
  What if we consider product attributes changed due to some regulation?
  How would consumers substitute towards or away from this product?
- **Shock 3: Merger** (usually modeled together with supply side for a firm level FOC)
  What if some brands merge, thereby changing the ownership matrix?
  How would equilibrium sales shift amongst all products?

Discrete Choice of Demand

Aggregate data vs. individual data

Aggregate data [Our focus for Topic 14]

- Our main data is product-level
- We observe for each product j, the price, the attribute, and sales (and ergo market share)
- Suppose you have 100 products and 5 years, ergo 500 observations
- The we eventually will regress market share (more on that later) to price and attributes in the example of multinomial logit
- Can enrich to allow random coefficient of individual heterogeneity with product-level data

Individual data

- Our main data is individual-level
- Suppose we observe 100 individuals, 5 choices in the choice set, 4 years of data
- We will have 20,000 obs
- For each individual, the choice not chosen will be coded as 0
- For each individual, the choice chosen will be coded as 1
- Run probabilistic regression of Logit, Probit, mixLogit, etc.

Discrete Choice of Demand

Energy economics and energy policy topics with aggregate data

Aggregate data:

- Our main data is product-level
- We observe for each product j, the price, the attribute, and sales (and ergo market share)

Market studied:

- Vehicle market
  What is the impact of subsidizing EVs or energy-efficient vehicles in general (a price shock)?
  What is the impact of the fuel economy standard that mandates vehicle fuel economy?
  What is the impact of subsidizing domestic products?
- Solar PV
  What is the impact of solar PV?
  How does the EV charging and solar PV complementarity affect this market?
- Appliance
  What is the impact of the Energy Star program?
- Fuel market (with station-level data)
- Consumer choice of retail electricity provider
- Consumer choice of retail natural gas provider
- Power plants' choice of coal supply from upstream suppliers
- Freight choice of fuel transportation

Discrete Choice of Demand

Energy economics and energy policy topics with individual data

Individual data:

- Our main data is individual-level
- Choices are coded as 0 and 1
- Run probability regression of Logit, Probit, mixLogit, etc.

Examples:

- Choice of travel mode: drive, carpool, public transit, bicycle, and walk w/ NHTS survey
- Enrollment of electricity pricing opt-in
- Cooling choice (the adoption of AC)
- Heating choice (heat-pump, gas furnace, etc)
- Choice of back-up electricity generator
- Choice of purchasing an energy-efficient home vs other counterparts
- Retrofit choice
- Solar PV adoption choice
- EV adoption
- Choice of scrapping vintage vehicles
- Power plants' choice amongst fuel types (to study fuel-switching)

Discrete Choice of Demand

Another related model

Discrete-Continuous models:

- Jointly estimate
  (i) the discrete choice of vehicle for the extensive margin of purchasing a vehicle +
  (ii) a continuous choice of VMT
- Example: Bento, Goulder, Jacobsen & von Haefen (2009) AER
  "Distributional and Efficiency Impacts of Increased US Gasoline Taxes"

Outline

- Introduction to Discrete Choice ✓
- Introduction to Simple Multinomial Discrete Choice with Aggregate Data
- Examples
- Key Practitioner Guide & Online Resource

Set-up

Define what is considered a market

Define a market $t \in \mathfrak{T}$

- Typical definition 1: national markets
  - It implies that each time period is a market
  - US vehicle market: a market t = a model year t (often Sep of calendar year $y-1$ to Aug in y)
  - Other vehicle market: a market t = a calendar year t
  - Other product market: a market t can be a month, a week, etc.
- Typical definition 2: regional markets
  - E.g., consider each country in the EU data in a year as a market
  - E.g., consider each state in the US data in a model year as a market
- Once we define a market, we can measure the total market size M_t
- With total market size, we can then measure the market share s_{jt} of a product j in market t

Set-up

Model individuals and their choices

- Each market has a set of individuals (consumers) $i \in \mathfrak{I}_i$
- Each market has a set of products $j \in \mathfrak{J}_t$ for individuals to choose from
  - $\mathfrak{J}_t \equiv [1, 2, \cdots, j, \cdots J_t]$ is known as the *(inner) choice set*
- In addition to choosing $j \in \mathfrak{J}_t$, individuals can choose the outside option $j = 0$
  - e.g., if your data has sales of all new vehicle purchases, outside option = buying a used car or no car
  - e.g., if your data has sales of all new & used vehicles purchased, outside option = buying no car
  - e.g., if your data is voting share, then outside option = no vote
- In multinomial logit, individuals do not have heterogeneous preferences, so why are we thinking about i?
  - We do think about i's discrete choice of purchasing j
  - We then rationalize consumer i's preference as a Logit probability
  - This lays a foundation of our market level analysis at the product level j when we translate individuals' choice probability into observed market share (a revealed preference approach)

Set-up

Measure market size M_t and market share s_{jt} for product j in market/year t

- Measure the total market size M_t given your definition of a market
  - Note: Choose M_t is an art. You should always refer to common practice in the market you study to construct the total market size
  - In the vehicle market, often ppl define the total market size using population
  - e.g., Grieco et al. (2024 QJE) defines the total market size M_t as the number of households in the US in year t divided by 2.5, as an approximation that avg hh owns about 2 vehicles and the average ownership time period is about 5 years
  - The size of M_t matters in counterfactual. The greater M_t, the more substitution towards outside good 0
- Then market share $s_{jt} = q_{jt}/M_{jt}$, with q being sales
- Outside option share s_{0t}
  - If we aggregate market share across all products, the remainder is s_{0t}
  - I.e., outside option share $s_{0t} = 1 - \sum_j s_{jt}$
- Eventually, in multinomial logit, the yvar is constructed from s_{jt} and s_{0t}
- Inner market share
  - We do not typically use this for estimation, but more for visualization
  - Inner market share is $q_{jt}/\sum_j q_{jt}$, so they add up to 1

Individual's Utility Maximization Problem

Assume no heterogeneous preference as in multinomial logit

$$\max_{j \in \mathfrak{J}_t \cup \{0\}} u_{ijt}, \text{ which we can define as }, u_{ijt} = \delta_{jt} + \varepsilon_{ijt}$$

- Element δ_{jt} is referred to as the mean utility
  - Note: Mean utility is at the product-level, ergo the name
  - We can parameterize a few things to enter the mean utility
  - E.g., in vehicle market, we can parametrize $\delta_{jt} = \alpha p_{jt} + \beta \mathbf{X}_{jt} + \phi_f + \phi_t + \eta_{jt}$
  - δ_{jt} includes price p (endogenous, more to that later)
  - δ_{jt} includes some product attributes (observed quality) X, and some FEs (ϕs)
  - δ_{jt} includes vehicle-specific unobservables (e.g., unobserved quality) η_{jt}
  - Q1: Why do we need this η_{jt}?
- Element ε_{ijt} is idiosyncratic taste
  - Unobservable consumer taste for a product
  - Q2: Why do we need this ε_{ijt}?
  - Assume ε_{ijt} follows an i.i.d. Gumbel/Type I extreme value distribution
  - A convenience assumption to have a nice, closed-form expression of choice probability

Forming choice probability of j above all else

We need to express $\mathbb{P}(u_{ijt} > u_{ikt}, \forall k \in \mathfrak{J}_t \cup \{0\})$

- Recall the UMP is $\max_j\{u_{ijt}\}$, where utilitiy $u_{ijt} = \delta_{jt} + \varepsilon_{ijt}$
- Then the choice probability of i picking j must be the likelihood for u_{ijt} to be greater than for all possible u_{ikt} for all other products $k \neq j$ and $k \in \mathfrak{J}_t \cup \{0\}$

$$\begin{aligned} \mathbb{P}(u_{ijt} > u_{ikt}, \forall k \neq j, k \in \mathfrak{J}_t \cup \{0\}) &= \mathbb{P}(\delta_{jt} + \varepsilon_{ijt} > \delta_{kt} + \varepsilon_{ikt}, \forall k \neq j, k \in \mathfrak{J}_t \cup \{0\}) \\ &= \mathbb{P}(\varepsilon_{ikt} < \delta_{jt} - \delta_{kt} + \varepsilon_{ijt}, \forall k \neq j, k \in \mathfrak{J}_t \cup \{0\}) \end{aligned} \tag{1}$$

Forming choice probability of j above all else

We need to express $\mathbb{P}(u_{ijt} > u_{ikt}, \forall k \in \mathfrak{J}_t \cup \{0\})$

- Recall the UMP is $\max_j\{u_{ijt}\}$, where utilitiy $u_{ijt} = \delta_{jt} + \varepsilon_{ijt}$
- Then the choice probability of i picking j must be the likelihood for u_{ijt} to be greater than for all possible u_{ikt} for all other products $k \neq j$ and $k \in \mathfrak{J}_t \cup \{0\}$

$$\begin{aligned} \mathbb{P}(u_{ijt} > u_{ikt}, \forall k \neq j, k \in \mathfrak{J}_t \cup \{0\}) &= \mathbb{P}(\delta_{jt} + \varepsilon_{ijt} > \delta_{kt} + \varepsilon_{ikt}, \forall k \neq j, k \in \mathfrak{J}_t \cup \{0\}) \\ &= \mathbb{P}(\varepsilon_{ikt} < \delta_{jt} - \delta_{kt} + \varepsilon_{ijt}, \forall k \neq j, k \in \mathfrak{J}_t \cup \{0\}) \end{aligned} \tag{1}$$

- Under Type 1 EV Distribution, PDF is $f(\varepsilon) = exp^{-\varepsilon} \cdot exp^{-exp^{-\varepsilon}}$
- Under Type 1 EV Distribution, CDF is $F(\varepsilon) = exp^{-exp^{-\varepsilon}}$
- Then the above choice probability transforms to a closed-form expression

Forming choice probability of j above all else

We need to express $\mathbb{P}(u_{ijt} > u_{ikt}, \forall k \in \mathfrak{J}_t \cup \{0\})$

- Recall the UMP is $\max_j \{u_{ijt}\}$, where utilitiy $u_{ijt} = \delta_{jt} + \varepsilon_{ijt}$
- Then the choice probability of i picking j must be the likelihood for u_{ijt} to be greater than for all possible u_{ikt} for all other products $k \neq j$ and $k \in \mathfrak{J}_t \cup \{0\}$

$$\begin{aligned} \mathbb{P}(u_{ijt} > u_{ikt}, \forall k \neq j, k \in \mathfrak{J}_t \cup \{0\}) &= \mathbb{P}(\delta_{jt} + \varepsilon_{ijt} > \delta_{kt} + \varepsilon_{ikt}, \forall k \neq j, k \in \mathfrak{J}_t \cup \{0\}) \\ &= \mathbb{P}(\varepsilon_{ikt} < \delta_{jt} - \delta_{kt} + \varepsilon_{ijt}, \forall k \neq j, k \in \mathfrak{J}_t \cup \{0\}) \end{aligned} \tag{1}$$

- Under Type 1 EV Distribution, PDF is $f(\varepsilon) = exp^{-\varepsilon} \cdot exp^{-exp^{-\varepsilon}}$
- Under Type 1 EV Distribution, CDF is $F(\varepsilon) = exp^{-exp^{-\varepsilon}}$
- Then the above choice probability transforms to a closed-form expression

$$\mathbb{P}(u_{ijt} > u_{ikt}, \forall k \neq j) = \frac{exp(\delta_{jt})}{\sum_{k \in \mathfrak{J}_t \cup \{0\}} \exp(\delta_{kt})} = \frac{exp(\delta_{jt})}{1 + \sum_{k \in \mathfrak{J}_t} \exp(\delta_{kt})} \tag{2}$$

- We normalize outside option to take utility $= 0$, i.e., $\delta_{0t} = 0$
- When there is no heterogeneity, market share s_{jt} becomes $\frac{\delta_{jt}}{1+\sum_k (\exp(\delta_{kt})}$

Derivation for Self-study: Details for Eq (2)

- Eq(1): $\mathbb{P}(u_{ijt} > u_{ikt},\ \forall k \neq j) = \mathbb{P}(\varepsilon_{ikt} < \delta_{jt} - \delta_{kt} + \varepsilon_{ijt},\ \forall k \neq j)$
- PDF: $f(\varepsilon) = e^{-\varepsilon} \cdot e^{-e^{-\varepsilon}}$
- CDF: $F(\varepsilon) = e^{-e^{-\varepsilon}}$
- I will suppress t for simplicity. For one $k \neq j$

$$\mathbb{P}(u_{ij} > u_{ik}) = \exp\{-\exp(-(\delta_j - \delta_k + \varepsilon_{ij}))\} = \exp\{-\exp(-\varepsilon_{ij}) \cdot \exp(\delta_k - \delta_j)\}$$

- Then for all $k \neq j$

$$\mathbb{P}(u_{ij} > u_{ik},\ \forall\, k \neq j) = \int \prod_{k \neq j} \exp\{-\exp(-\varepsilon_{ij}) \cdot \exp(\delta_k - \delta_j)\}\, f(\varepsilon_{ij})\, d\varepsilon_{ij}, \quad \text{Note: } \prod_k \exp(x_k) = \exp\left(\sum_k x_k\right)$$

$$= \int \exp\left\{\sum_{k \neq j} \left[-\exp(-\varepsilon_{ij}) \cdot \exp(\delta_k - \delta_j)\right]\right\} f(\varepsilon_{ij})\, d\varepsilon_{ij}$$

$$= \int \exp\left\{-\exp(-\varepsilon_{ij}) \cdot \sum_{k \neq j} \exp(\delta_k - \delta_j)\right\} f(\varepsilon_{ij})\, d\varepsilon_{ij}\,, \quad \text{Denote the invariant part as } A$$

$$= \int \exp\left\{-\exp(-\varepsilon_{ij}) \cdot A\right\} f(\varepsilon_{ij})\, d\varepsilon_{ij}$$

Derivation for Self-study: Details for Eq (2)

- PDF: $f(\varepsilon) = e^{-\varepsilon} \cdot e^{-e^{-\varepsilon}}$
- Denote $\sum_{k \neq j} \exp(\delta_k - \delta_j)$ as A for now

$$
\begin{aligned}
\mathbb{P}(u_{ij} > u_{ik},\ \forall\, k \neq j) &= \int_{-\infty}^{+\infty} \exp\left\{-\exp(-\varepsilon_{ij}) \cdot A\right\}\ f(\varepsilon_{ij})\, d\varepsilon_{ij} \\
&= \int_{-\infty}^{+\infty} \exp\left\{-\exp(-\varepsilon_{ij}) \cdot A\right\} \cdot \exp(-\varepsilon_{ij}) \cdot \exp\left\{-\exp(-\varepsilon_{ij})\right\}\ d\varepsilon_{ij} \\
&= \int_{-\infty}^{+\infty} \exp(-\varepsilon_{ij}) \cdot \exp\left\{-\exp(-\varepsilon_{ij}) \cdot A\right\} \cdot \exp\left\{-\exp(-\varepsilon_{ij})\right\}\ d\varepsilon_{ij} \\
&= \int_{-\infty}^{+\infty} \exp(-\varepsilon_{ij}) \cdot \exp\left\{-\exp(-\varepsilon_{ij}) \cdot (A+1)\right\}\ d\varepsilon_{ij}
\end{aligned}
$$

- Denote $\nu = e^{-\varepsilon}$

$$
\begin{aligned}
\mathbb{P}(u_{ij} > u_{ik},\ \forall\, k \neq j) &= \int_{-\infty}^{+\infty} \nu \cdot \exp\{\nu \cdot (A+1)\}\ d\varepsilon_{ij}\text{ , next consider } d\nu/d\varepsilon = -\nu \\
&= \int_{+\infty}^{0} \exp\{-\nu(A+1)\}\,(-d\nu) = \int_{0}^{+\infty} \exp\{-\nu(A+1)\}\, d\nu
\end{aligned}
$$

Derivation for Self-study: Details for Eq (2)

$$
\begin{aligned}
\mathbb{P}(u_{ij} > u_{ik}, \ \forall\, k \neq j) &= \int_0^{+\infty} \exp\left\{-\nu(A+1)\right\} d\nu \\
&= \left[-\frac{1}{A+1}\exp\{-\nu(A+1)\}\right]_0^{+\infty} = -\frac{1}{A+1}\left[\exp\{-\nu(A+1)\}\right]_0^{+\infty} \\
&= -\frac{1}{A+1}\left\{\left[\exp\{-\nu(A+1)\}\right]_{\nu\to+\infty} - \left[\exp\{-\nu(A+1)\}\right]_{\nu=0}\right\} = -\frac{1}{A+1}\{0-1\} = \frac{1}{A+1} \\
&= \frac{1}{1+\sum_{k\neq j}\exp(\delta_k-\delta_j)}, \text{ now recall } A \equiv \sum_{k\neq j}\exp(\delta_k-\delta_j) \\
&= \frac{1}{1+\sum_{k\neq j}\exp(\delta_k)\cdot\exp(-\delta_j))} = \frac{\exp(\delta_j)}{\exp(\delta_j)+\sum_{k\neq j}\exp(\delta_k)} \\
&= \frac{exp(\delta_j)}{\sum_{k\in\mathfrak{J}\cup\{0\}}\exp(\delta_k)} = \frac{exp(\delta_j)}{1+\sum_{k\in\mathfrak{J}}\exp(\delta_k)}
\end{aligned}
$$

- Translate choice probability into market share s_{jt}

$$
s_{jt} = \frac{exp(\delta_{jt})}{1+\sum_{k\in\mathfrak{J}_t}\exp(\delta_{kt})} \tag{3}
$$

Revealed-preference Approach to Rationalize Market Share s_{jt}

$$s_{jt} = \frac{exp(\delta_{jt})}{1 + \sum_{k \in \mathfrak{J}_t} \exp(\delta_{kt})} \tag{3}$$

- Note: This is already market level
- Recall that we use observables to parameterize mean utility $\delta_{jt} = \alpha p_{jt} + \beta \mathbf{X}_{jt} + \phi_f + \phi_t + \eta_{jt}$
- This means we can fit data at the product level to match market share
- Price coefficient α will tell you about price sensitivity, β will tell you tastes for attributes
- The goal of estimation is to estimate the price coefficient (later elasticity) and other taste parameters to rationalize the observed market share (reveal preference!)
- Now we actually do not estimate Eq (3) per se, but a much simpler version

Forming Market Share Regression

- Market share for j: $$s_{jt} = \frac{\exp(\delta_{jt})}{1+\sum_{k\in\mathfrak{J}_t}\exp(\delta_{kt})}$$
- Market share for outside option $j=0$: $s_{0t} = \frac{\exp(0)}{1+\sum_{k\in\mathfrak{J}_t}\exp(\delta_{kt})} = \frac{1}{1+\sum_{k\in\mathfrak{J}_t}\exp(\delta_{kt})}$
- Relative market share: $\frac{s_{jt}}{s_{0t}} = \exp(\delta_{jt})$
- Take a log transformation

$$\ln\left(\frac{s_{jt}}{s_{0t}}\right) = \delta_{jt}$$
$$\ln\left(\frac{s_{jt}}{s_{0t}}\right) = \alpha p_{jt} + \beta \mathbf{X}_{jt} + \phi_f + \phi_t + \eta_{jt} \tag{4}$$

- It is simply a linear regression! Implicitly under the assumption $\eta_{jt} \sim \textit{Normal}$
- Hence the name of our topic **simple multinomial logit**
- Although mathematically the LHS can also be written as $\ln s_{jt} - \ln s_{0t}$, I recommend computing $\ln\left(\frac{s_{jt}}{s_{0t}}\right)$ in case computer store small number incorrectly

Multinomial Demand Estimation

Issues: (1) What can go wrong? (2) What can we use estimates for?

For product j in market t, we estimate

$$\ln\left(\frac{s_{jt}}{s_{0t}}\right) = \alpha p_j + \mathbf{X}_j\beta + \phi_f + \phi_t + \eta_{jt} \tag{4}$$

[Next] Challenge

- Often, we care about the price coefficient, as it eventually will lead us to demand elasticity
- That means we want to have a reliable estimate $\hat{\alpha}$
- But price is endogenous. Why?
- Solution: IV(s)
- All linear regression commands with IV feature can apply: `reghdfe` in Stata; `feols` in R; `linearmodels.iv` in `Python`. Also, can use non-random-coefficient version of `PyBLP`

[Next] What we use estimates for

- Obtain own- and cross-elasticity matrix and other outputs
- Use estimates to simulate counterfactuals

Outline

- Introduction to Discrete Choice ✓
- Introduction to Simple Multinomial Discrete Choice with Aggregate Data ✓
- Output from Demand Estimation: Elasticities
- IV
- Key Practitioner Guide & Online Resource

Typical Outputs from Demand Estimation

1. Elasticities

For product j in market t, suppose we estimate a simple multinomial logit

$$\ln\left(\frac{s_{jt}}{s_{0t}}\right) = \alpha p_{jt} + \mathbf{X}_{jt}\beta + \phi_f + \phi_t + \eta_{jt} \tag{4}$$

- Point estimates of α and β can translate into elasticities

$$\text{Elasticity (own-price) } \varepsilon^d_{jj} = \alpha \cdot p_{jt} \cdot (1 - s_{jt})$$
$$\text{Elasticity (cross-price) } \varepsilon^d_{jk} = -\alpha \cdot p_{kt} \cdot s_{kt}$$

- These are essentially the diagonal and off-diagonal elements of an elasticity matrix

Typical Outputs from Demand Estimation

1. Elasticities

For product j in market t, suppose we estimate a simple multinomial logit

$$\ln\left(\frac{s_{jt}}{s_{0t}}\right) = \alpha p_{jt} + \mathbf{X}_{jt}\beta + \phi_f + \phi_t + \eta_{jt} \tag{4}$$

- Point estimates of α and β can translate into elasticities

$$\text{Elasticity (own-price) } \varepsilon_{jj}^{d} = \alpha \cdot p_{jt} \cdot (1 - s_{jt})$$

$$\text{Elasticity (cross-price) } \varepsilon_{jk}^{d} = -\alpha \cdot p_{kt} \cdot s_{kt}$$

- These are essentially the diagonal and off-diagonal elements of an elasticity matrix
- We can also describe elasticity regarding an attribute
- Suppose there is only one attribute x in $\mathbf{X}$ for simplicity

$$\text{Elasticity (own-attribute x) } \varepsilon_{jj}^{x} = \beta \cdot x_{jt} \cdot (1 - s_{jt})$$

$$\text{Elasticity (cross-attribute x) } \varepsilon_{jk}^{x} = -\beta \cdot x_{jt} \cdot s_{jt}$$

- These assume price and attribute enter linearly
- If you use $\ln p_{jt}$, own price elasticity is $\alpha \cdot (1 - s_{jt})$

Typical Outputs from Demand Estimation

1. Elasticities (Derive)

- Recall market share for j: $s_{jt} = \frac{\exp(\delta_{jt})}{1+\sum_{k\in\mathfrak{J}_t}\exp(\delta_{kt})}$
- To derive elasticity, we should first derive the slope ds/dp
- Take price as an example
- Step 1: we will describe $\frac{\partial s_{jt}}{\partial p_{jt}}$ and $\frac{\partial s_{jt}}{\partial p_{kt}}$ (this forms response matrix Λ)
- Step 2: we will then normalize the response by p and s to express elasticity ε^p

$$\Lambda^p_{jk} = \frac{\partial s_j}{\partial p_k} = \begin{cases} \frac{\partial s_j}{\partial p_j} & \text{if } k = j \\ \frac{\partial s_j}{\partial p_k} & \text{if } k \neq j \end{cases}$$

$$\varepsilon^p_{jk} = \frac{\partial s_j}{\partial p_k}\cdot\frac{p_k}{s_j} = \begin{cases} \frac{\partial s_j}{\partial p_j}\cdot\frac{p_j}{s_j} & \text{if } k = j \\ \frac{\partial s_j}{\partial p_k}\cdot\frac{p_k}{s_j} & \text{if } k \neq j \end{cases}$$

$$\Lambda^p = \begin{bmatrix} \frac{\partial s_1}{\partial p_1} & \frac{\partial s_1}{\partial p_2} & \cdots & \frac{\partial s_1}{\partial p_J} \\ \frac{\partial s_2}{\partial p_1} & \frac{\partial s_2}{\partial p_2} & \cdots & \frac{\partial s_2}{\partial p_J} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial s_J}{\partial p_1} & \frac{\partial s_J}{\partial p_2} & \cdots & \frac{\partial s_J}{\partial p_J} \end{bmatrix}_{\text{J X J}}$$

$$\varepsilon^p = \begin{bmatrix} \frac{\partial s_1}{\partial p_1}\cdot\frac{p_1}{s_1} & \frac{\partial s_1}{\partial p_2}\cdot\frac{p_2}{s_1} & \cdots & \frac{\partial s_1}{\partial p_J}\cdot\frac{p_J}{s_1} \\ \frac{\partial s_2}{\partial p_1}\cdot\frac{p_1}{s_2} & \frac{\partial s_2}{\partial p_2}\cdot\frac{p_2}{s_2} & \cdots & \frac{\partial s_2}{\partial p_J}\cdot\frac{p_J}{s_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial s_J}{\partial p_1}\cdot\frac{p_1}{s_J} & \frac{\partial s_J}{\partial p_2}\cdot\frac{p_2}{s_J} & \cdots & \frac{\partial s_J}{\partial p_J}\cdot\frac{p_J}{s_J} \end{bmatrix}_{\text{J X J}}$$

Typical Outputs from Demand Estimation

1. Elasticities (Derive diagonal own-effect terms)

- Recall market share for j: $s_{jt} = \frac{\exp(\delta_{jt})}{1+\sum_{k\in\mathfrak{J}_t}\exp(\delta_{kt})}$
- Mean utility $\alpha p_{jt} + \mathbf{X}_{jt}\beta + \phi_f + \phi_t + \eta_{jt}$
- Denote the denominator for market share as D_t for simplicity: $s_{jt} = \frac{\exp(\delta_{jt})}{D_t}$
- Then own-price response Λ^p_{jj} and own-price elasticity ε^d_{jj}

$$\Lambda^p_{jj} = \frac{\partial s_{jt}}{\partial p_{jt}} = \alpha\exp(\delta_{jt})\cdot\frac{1}{D_t} + \exp(\delta_{jt})\cdot\left[-\frac{1}{D_t^2}\cdot\alpha\exp(\delta_{jt})\right]$$

$$= \frac{\alpha\exp(\delta_{jt})\cdot D_t - \alpha[\exp(\delta_{jt})]^2}{D_t^2} = \alpha\cdot\underbrace{\frac{\exp(\delta_{jt})}{D_t}}_{=s_{jt}}\cdot\underbrace{\left(1-\frac{\exp(\delta_{jt})}{D_t}\right)}_{=1-s_{jt}} = \alpha\cdot s_{jt}(1-s_{jt})$$

$$\varepsilon^d_{jj} = \frac{\partial s_{jt}}{\partial p_{jt}}\cdot\frac{p_{jt}}{s_{jt}} = \alpha\cdot s_{jt}(1-s_{jt})\cdot\frac{p_{jt}}{s_{jt}} = \alpha\cdot p_{jt}\cdot(1-s_{jt})$$

Typical Outputs from Demand Estimation

1. Elasticities (Derive off-diagonal cross-effect terms)

- Recall market share for j: $s_{jt} = \frac{\exp(\delta_{jt})}{1+\sum_{k\in\mathfrak{J}_t}\exp(\delta_{kt})}$
- Mean utility $\alpha p_{jt} + \mathbf{X}_{jt}\beta + \phi_f + \phi_t + \eta_{jt}$
- Denote the denominator for market share as D_t for simplicity: $s_{jt} = \frac{exp(\delta_{jt})}{D_t}$
- Then cross-price response Λ^p_{jk} and cross-price elasticity ε^d_{jk}

$$\Lambda^p_{jk} = \frac{\partial s_{jt}}{\partial p_{kt}} = 0\cdot\frac{1}{D_t} + \exp(\delta_{jt})\cdot\left[-\frac{1}{D_t^2}\cdot\alpha\exp(\delta_{kt})\right] = -\frac{\alpha\exp(\delta_{jt})\exp(\delta_{kt})}{D_t^2}$$

$$= -\alpha\cdot\underbrace{\frac{\exp(\delta_{jt})}{D_t}}_{=s_{jt}}\cdot\underbrace{\frac{\exp(\delta_{kt})}{D_t}}_{=s_{kt}} = -\alpha\cdot s_{jt}s_{kt}$$

$$\varepsilon^d_{jk} = \frac{\partial s_{jt}}{\partial p_{kt}}\cdot\frac{p_{kt}}{s_{jt}} = -\alpha\cdot s_{jt}s_{kt}\cdot\frac{p_{kt}}{s_{jt}} = -\alpha\cdot p_{kt}\cdot s_{kt}$$

Typical Outputs from Demand Estimation

1. Elasticities (Summary)

- **Response matrix** (suppress t for simplicity)

$$\text{Individual terms}: \Lambda^p_{jk} = \frac{\partial s_j}{\partial p_k} = \begin{cases} \frac{\partial s_j}{\partial p_j} = \alpha s_j(1 - s_j) & \text{if } k = j \\ \frac{\partial s_j}{\partial p_k} = \alpha s_j s_k & \text{if } k \neq j \end{cases}$$

$$\text{Matrix version}: \Lambda^p = \begin{bmatrix} \alpha s_1(1 - s_1) & -\alpha s_1 s_2 & \cdots & -\alpha s_1 s_J \\ -\alpha s_2 s_1 & \alpha s_2(1 - s_2) & \cdots & -\alpha s_2 s_J \\ \vdots & \vdots & \ddots & \vdots \\ -\alpha s_J s_1 & -\alpha s_J s_2 & \cdots & \alpha s_{iJ}(1 - s_J) \end{bmatrix}_{J \times J}$$

- **Elasticity matrix** (suppress t for simplicity)

$$\text{Individual terms}: \varepsilon^p_{jk} = \frac{\partial s_j}{\partial p_k} \cdot \frac{p_k}{s_j} = \begin{cases} \frac{\partial s_j}{\partial p_j} \cdot \frac{p_j}{s_j} = \alpha p_j(1 - s_j) & \text{if } k = j \\ \frac{\partial s_j}{\partial p_k} \cdot \frac{p_k}{s_j} = -\alpha p_k s_k & \text{if } k \neq j \end{cases}$$

$$\text{Matrix version}: \boldsymbol{\varepsilon^p} = \begin{bmatrix} \alpha(1 - s_1)p_1 & -\alpha s_2 p_2 & \cdots & -\alpha s_J p_J \\ -\alpha s_1 p_1 & \alpha(1 - s_2)p_2 & \cdots & -\alpha s_J p_J \\ \vdots & \vdots & \ddots & \vdots \\ -\alpha s_1 p_1 & -\alpha s_2 p_2 & \cdots & \alpha(1 - s_J)p_J \end{bmatrix}_{J \times J}$$

Typical Outputs from Demand Estimation

1. Elasticities and Limitations of Multinomial Logit

$$\varepsilon_{jk}^{p} = \frac{\partial s_j}{\partial p_k} \cdot \frac{p_k}{s_j} = \begin{cases} \frac{\partial s_j}{\partial p_j} \cdot \frac{p_j}{s_j} = \alpha p_j (1 - s_j) & \text{if } k = j \\ \frac{\partial s_j}{\partial p_k} \cdot \frac{p_k}{s_j} = -\alpha p_k s_k & \text{if } k \neq j \end{cases}$$

Price sensitivity

- Own-price elasticity is an increase in p_j! It is also declining in share s_j!
- Why is it bad?
- Higher price product will have elasticity, and ergo, lower mark-up
- Extremely \$\$\$ product will have extreme low mark-up
- Extremely popular product with high s_j will have low elasticity, and ergo, higher markup
- Is this real?

Typical Outputs from Demand Estimation

1. Elasticities and Limitations of Multinomial Logit

$$\varepsilon_{jk}^{p} = \frac{\partial s_j}{\partial p_k} \cdot \frac{p_k}{s_j} = \begin{cases} \frac{\partial s_j}{\partial p_j} \cdot \frac{p_j}{s_j} = \alpha p_j (1 - s_j) & \text{if } k = j \\ \frac{\partial s_j}{\partial p_k} \cdot \frac{p_k}{s_j} = -\alpha p_k s_k & \text{if } k \neq j \end{cases}$$

Substitution

- Cross-price elasticity is decreasing in p_k! It is also declining in share s_j!
- Why is it bad? What makes substitution high? Is that real?
- The issue of proportional substitution
  - Example 1. Imagine if p_k drops, then all other products' substitution towards k in percentage terms is proportional to market share s_k
  - Example 2. Imagine if p_j drops, then for any other product k and l, $\frac{\varepsilon_{jk}}{\varepsilon_{jl}} = \frac{p_k s_k}{p_l s_l}$
    The relative cross-price elasticity is fixed, and doesn't depend on which of k and l is a closer rival to j
    This is similar to the blue and red bus problem
  - Example 3. An extreme case of Example 2
    Consider k and l, and k is a closer substitute for j than l
    Suppose while k and l are different, they have identical price and market share, then $\frac{\varepsilon_{jk}}{\varepsilon_{jl}} = 1$!!
    It means when j gets expensive, the change to divert to k and l (in percentage terms) are identical!
  - Examples break Independence of Irrelevant Alternatives (IIA)

Typical Outputs from Demand Estimation

2. WTP for attribute

For product j in market t, suppose we estimate a simple multinomial logit

$$\ln\left(\frac{s_{jt}}{s_{0t}}\right) = \alpha p_{jt} + \mathbf{X}_{jt}\beta + \phi_f + \phi_t + \eta_{jt} \tag{4}$$

- Point estimates of α and β can translate into WTP
- Suppose there is only one attribute x in vector $\mathbf{X}$ for simplicity
- Suppose x is mpg, fuel economy of the vehicle
- Suppres t for simplicity

$$\text{Response (own-price)} \quad \Lambda^p_{jj} = \frac{\partial s_j}{\partial p_j} = \alpha \cdot p_j \cdot (1 - s_j)$$

$$\text{Response (own-attribute x)} \quad \Lambda^x_{jj} = \frac{\partial s_j}{\partial x_j} = \beta \cdot x_j \cdot (1 - s_j)$$

- Then **Marginal WTP for an increase in attribute x**

$$MWTP_x = \frac{\partial p_j}{\partial x_j} = \frac{\beta}{\alpha} \tag{5}$$

Typical Outputs from Demand Estimation

3. Diversion Ratio

- Recall the elements in the response matrix

$$\Lambda^p_{jk} = \frac{\partial s_j}{\partial p_k} = \begin{cases} \frac{\partial s_j}{\partial p_j} = \alpha s_j(1-s_j) & \text{if } k = j \\ \frac{\partial s_j}{\partial p_k} = \alpha s_j s_k & \text{if } k \neq j \end{cases}$$

- **Diversion ratio** is the relative response

$$\mathbf{D_{j\to k}} = \frac{\frac{\partial s_k}{\partial p_j}}{\left|\frac{\partial s_j}{\partial p_j}\right|} = \begin{cases} 1 & \text{if } k = j \\ \frac{s_k}{1-s_j} & \text{if } k \neq j \end{cases}, \quad \text{towards outside option } \mathbf{D_{j\to 0}} = \frac{s_0}{1-s_j}$$

- Multinomial means a constant diversion ratio from j to k, if there is a shock to another product l (e.g., price change, removal, entry)
- It breaks IIA

Typical Outputs from Demand Estimation

Summary of three typical outputs

- **1. Elasticities** (use price as an example (suppress t for simplicity)

$$\text{Individual terms}: \varepsilon^{p}_{jk} = \frac{\partial s_j}{\partial p_k} \cdot \frac{p_k}{s_j} = \begin{cases} \frac{\partial s_j}{\partial p_j} \cdot \frac{p_j}{s_j} = \alpha p_j(1 - s_j) & \text{if } k = j \\ \frac{\partial s_j}{\partial p_k} \cdot \frac{p_k}{s_j} = -\alpha p_k s_k & \text{if } k \neq j \end{cases}$$

- **2. Marginal WTP for an increase in attribute x**

$$MWTP_x = \frac{\partial p_j}{\partial x_j} = \frac{\beta}{\alpha}$$

- **3. Diversion ratios**

$$\mathbf{D_{j\to k}} = \frac{\frac{\partial s_k}{\partial p_j}}{\left|\frac{\partial s_j}{\partial p_j}\right|} = \begin{cases} 1 & \text{if } k = j \\ \frac{s_k}{1-s_j} & \text{if } k \neq j \end{cases}, \quad \text{towards outside option } \mathbf{D_{j\to 0}} = \frac{s_0}{1 - s_j}$$

Outline

- Introduction to Discrete Choice ✓
- Introduction to Simple Multinomial Discrete Choice with Aggregate Data ✓
- Output from Demand Estimation: Elasticities ✓
- IV
- Key Practitioner Guide & Online Resource

Endogeneity Issue for the Price

Common culprits and confounders

$$\ln\left(\frac{s_{jt}}{s_{0t}}\right) = \alpha p_j + \mathbf{X}_j\beta + \phi_f + \phi_t + \eta_{jt} \quad (4)$$

$$\delta_{jt} = \alpha p_j + \mathbf{X}_j\beta + \phi_f + \phi_t + \eta_{jt} \quad (6)$$

- Often, we care about the price coefficient, as it eventually will lead us to demand elasticity
- That means we want to have a reliable estimate $\hat{\alpha}$
- But price is endogenous. Why? What is potentially in η_{jt}?

Endogeneity Issue for the Price

Common culprits and confounders

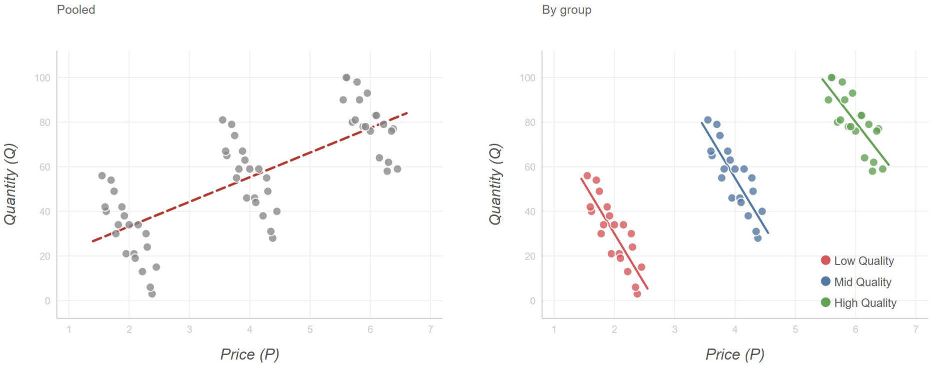
$$\ln\left(\frac{s_{jt}}{s_{0t}}\right) = \alpha p_j + \mathbf{X}_j\beta + \phi_f + \phi_t + \eta_{jt} \tag{4}$$

$$\delta_{jt} = \alpha p_j + \mathbf{X}_j\beta + \phi_f + \phi_t + \eta_{jt} \tag{6}$$

- Often, we care about the price coefficient, as it eventually will lead us to demand elasticity
- That means we want to have a reliable estimate $\hat{\alpha}$
- But price is endogenous. Why? What is potentially in η_{jt}?
- **Unobserved quality**
  - Let's call it κ_{jt}. Suppose we are talking about desirable qualities
  - Then $corr(p, \kappa) > 0$ and $corr(\delta, \kappa) > 0$, so that $bias > 0$
  - The true price coefficient: $\underset{(?)}{\alpha_{naive}} = \underset{(+)}{\alpha_{true}} + \underset{(+)}{bias}$
  - We may estimate too small a negative coefficient, i.e., less elastic than the true elasticity
  - In extreme cases, when the bias term is too sizable, we may incorrectly detect an upward-sloping demand curve!

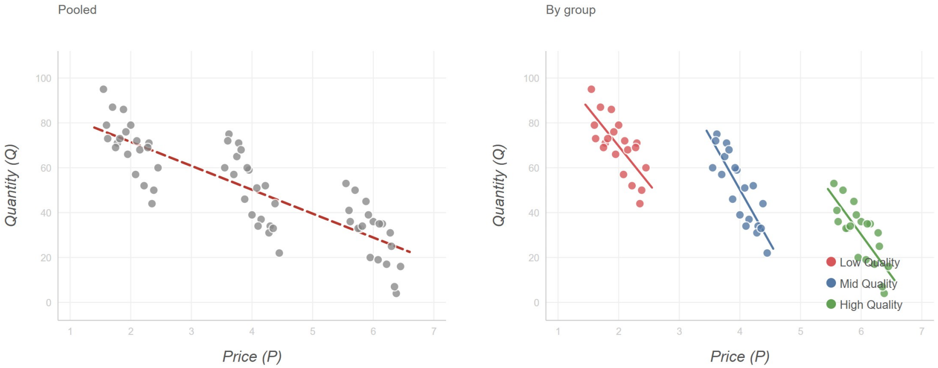
Endogeneity from Unobserved Quality

Example 1: Naive positive demand curve



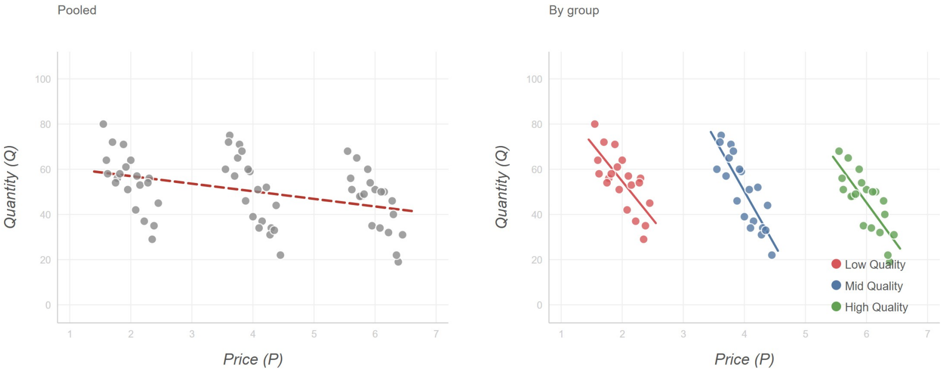
Endogeneity from Unobserved Quality

Example 2: We underestimate demand elasticity



Endogeneity from Unobserved Quality

Example 3: We underestimate demand elasticity so much so it is almost zero



Endogeneity Issue for the Price

Common culprits and confounders

$$\ln\left(\frac{s_{jt}}{s_{0t}}\right) = \alpha p_j + \mathbf{X}_j\beta + \phi_f + \phi_t + \eta_{jt} \quad (4)$$

$$\delta_{jt} = \alpha p_j + \mathbf{X}_j\beta + \phi_f + \phi_t + \eta_{jt} \quad (6)$$

- **Unobserved quality**
  - Here, *bias* > 0 for some products so we may overestimate the negative coef
- **Common unobserved demand shocks**
  - E.g., a booming economy that increases both P and Q
    $\rightarrow$ Here, *bias* > 0 so we may underestimate the negative coef
  - E.g., a wave of baby boom so both P and Q go up for large vehicles
    $\rightarrow$ Here, *bias* > 0 for some products so we may underestimate the negative coef
  - E.g., severe oil price shock may increase P and reduce Q
    $\rightarrow$ Here, *bias* < 0 for some products so we may overestimate the negative coef
- The most typical case is that the naive OLS will bias us towards zero

Endogeneity Issue for the Price

Solutions

$$\ln\left(\frac{s_{jt}}{s_{0t}}\right) = \alpha p_j + \mathbf{X}_j\beta + \phi_f + \phi_t + \eta_{jt} \tag{4}$$

$$\delta_{jt} = \alpha p_j + \mathbf{X}_j\beta + \phi_f + \phi_t + \eta_{jt} \tag{6}$$

- Panel data structure and FE
  - When data is national-level, can include year FE, product segment FE
  - When data is regional-level, can include year FE, product segment FE, location FE
  - Doing so will allow us to do some "partialing-out"
  - If you add year FE ϕ_t, you implicitly allow the mean utility of the outside option δ_{0t} to vary by year, absorbing aggregate demand shocks common to all products in market t
- IVs:
  - We will use a GMM estimator
  - Implementation is super easy! All IV regression packages can apply:
    e.g., `reghdfe` in Stata;
    e.g., `feols` in R;
    e.g., `linearmodels.iv` in `Python`. Also, the non-random-coefficient version of `PyBLP`

Instrumental Variables

$$\ln\left(\frac{s_{jt}}{s_{0t}}\right) = \alpha p_j + \mathbf{X}_j\beta + \phi_f + \phi_t + \eta_{jt} \tag{4}$$

$$\delta_{jt} = \alpha p_j + \mathbf{X}_j\beta + \phi_f + \phi_t + \eta_{jt} \tag{6}$$

- Types of IVs
  - Markup shifters
  - Cost shifters
  - How would they work?
  - Or a combination of both, e.g., introduce both or make some interactions

Instrumental Variables

$$\ln\left(\frac{s_{jt}}{s_{0t}}\right) = \alpha p_j + \mathbf{X}_j\beta + \phi_f + \phi_t + \eta_{jt} \quad (4)$$

$$\delta_{jt} = \alpha p_j + \mathbf{X}_j\beta + \phi_f + \phi_t + \eta_{jt} \quad (6)$$

- Types of IVs
  - Markup shifters
  - Cost shifters
  - How would they work?
  - Or a combination of both, e.g., introduce both or make some interactions
- Typical styles of shifters
  - BLP-style IVs (and Train-style IVs)
  - Hausman-style IVs
  - Waldfogel-style IVs
  - Recent work:
    - Fan-style IV (a refined BLP IV that combined Hausman and Waldfogel's ideas)
    - Optimal IV (a huge literature, e.g., Reynaert & Verboven 2013)
    - Differentiation IV (Gandhi & Houde 2020)

IV: Markup Shifters

The most classic IVs: The BLP IVs

BLP IVs include three parts: (i) number of other products on the market, (ii) exogenous attributes from rival firms, (iii) exogenous attributes from the same firm

$$\underbrace{z_{jt}^{\text{num}} = |\mathcal{J}_t|}_{\text{no. of products in } t} \qquad \underbrace{z_{jt}^{\text{own}} = \sum_{k \in \mathcal{J}_{ft},\, k \neq j} \mathbf{z}_{kt}}_{\text{other products, same firm } f} \qquad \underbrace{z_{jt}^{\text{rival}} = \sum_{k \in \mathcal{J}_{-f,t}} \mathbf{z}_{kt}}_{\text{products by rival firms}}$$

- How do BLP IVs work?

IV: Markup Shifters

The most classic IVs: The BLP IVs

BLP IVs include three parts: (i) number of other products on the market, (ii) exogenous attributes from rival firms, (iii) exogenous attributes from the same firm

$$z_{jt}^{\text{num}} = |\mathcal{J}_t| \qquad z_{jt}^{\text{own}} = \sum_{k \in \mathcal{J}_{ft},\, k \neq j} \mathbf{Z}_{kt} \qquad z_{jt}^{\text{rival}} = \sum_{k \in \mathcal{J}_{-f,t}} \mathbf{Z}_{kt}$$

no. of products in t — other products, same firm f — products by rival firms

- How do BLP IVs work?
- The exogenous attributes are attributes that are not already included in the mean utility
- What are typical "exogenous attributes"?
  Why are they exogenous? Are they exogenous at all?
  What are better choices of attributes?
- What are potential issues with BLP IVs?
  - 1: Potential bias
  - 2: Potential weakness if the number of products is too large (Armstrong 2016). Why?

IV: Markup Shifters

The most classic IVs: The BLP IVs

BLP IVs include three parts: (i) number of other products on the market, (ii) exogenous attributes from rival firms, (iii) exogenous attributes from the same firm

$$z_{jt}^{\text{num}} = |\mathcal{J}_t|$$
no. of products in t

$$z_{jt}^{\text{own}} = \sum_{k \in \mathcal{J}_{ft},\, k \neq j} \mathbf{z}_{kt}$$
other products, same firm f

$$z_{jt}^{\text{rival}} = \sum_{k \in \mathcal{J}_{-f,t}} \mathbf{z}_{kt}$$
products by rival firms

- How do BLP IVs work?
- The exogenous attributes are attributes that are not already included in the mean utility
- What are typical "exogenous attributes"?
  Why are they exogenous? Are they exogenous at all?
  What are better choices of attributes?
- What are potential issues with BLP IVs?
  - 1: Potential bias
  - 2: Potential weakness if the number of products is too large (Armstrong 2016). Why?
- Derivation of BLP-IV: Train's idea
  Calculate the distance between z_{jt} to $\overline{z_{-j,t}}$

IV: Markup Shifters

Waldfogel IVs: Demographic shifters

Waldfogel suggest to use demographics shifters as IVs

- This IV works the best if you have regional data
- So a market = a location X a period (e.g., state by year)
- Miravete et al. (2022 Ecta) IV = share of young consumers in each market to IV for vodka price
- Note that this IV doesn't vary by j and is constant within a market, so it may be weak
- However, if you have a consumer-level survey, you can also compute product-level demographics as IV
- It does inspire Fan-style IV using demographics of nearby market (coming next)
- But let's discuss cost shifters first

IV: Cost Shifters

Hausman IVs

Hausman suggests that if you have regional data, then use you can use product j's price in another location to IV for product j's price in current location

- Example: Nevo (2021) price of cereal in another city
- How does Hausman IV work?
- Denote current location as $r1$, and another location as $r2$
- $p_j^{r2} = mc_j + markup_j^{r2}$
- Typically, we can assume each market has its own demand shock
- But cost shocks across multiple locations (markets) are correlated
  e.g., the supply-chain shock of BMW will affect all states
- So p_j^{r2} and p_j^{r1} will correlate due to the common MC!
- While the markup part in $r2$ is unrelated to location $r1$ as demand shock can be independent
- What's the catch? Often, not every one have the luxury to have regional-level sales data together with regional-level price variation

IV: Cost Shifters in General

Very tricky to find...

- If you have a cost shifter that varies at the product level, that will be ideal
  - E.g., what would make vehicle j more expensive to produce than vehicle k from one year to the next?
  - Material cost: e.g., some cars need one type of chip, and another uses a different chip, and then chip prices vary across products and over time
  - Labor cost: can be a potential, but (i) we need to know the production location and (ii) the labor market needs to have a lot of variation
  - Or shipping cost
  - Catch: How realistic can we observe this type of data? Also, how much variation is there? Is the IV going to be weak?
- It is really difficult to find!
- Special cases: Grieco et al. (2024)
  - Each vehicle has its own primary production location
  - Each location has a different exchange rate

IV: Cost Shifters

Use Hausman and Waldfogel together

A derivation version of Waldfogel IV is to use the nearby market's demographics

- This combines both Woldfogel and Hausman
- It uses Waldfogel's markup-shifter idea of demographics
- It uses Hausman's cost-shifter idea of nearby/other location/market
- Think location $r1$ and $r2$ again
- $p_j^{r2} = mc_j + markup_j^{r2}$
- Demographics in $r2$ is a markup shifter for $r2$
- However, if price of luxury products in $r2$ increase b/c income rise in $r2$, it may create some economies of scale for the firm to serve both markets, so that mc will be affected
- Then indirectly price in $r1$ will be affected as mc can be shifted due to demographics in $r2$

IV: Markup Shifters

Fan-style IV: A refined BLP-IV with some Waldfogel idea

- This IV uses rival product k's demographics in nearby market
- This IV requires us not only to observe demographics at the market level, but also at the product level

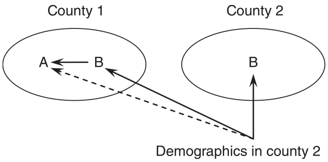


FIGURE 2. INSTRUMENTS

- Newspaper prices vary by year but not by location
- Consumer demographics for newspaper B in county 2 is markup shifter for product B in $r2$ affecting p_B^{r2}
- Given pricing is the same between locations, it will affect p_B^{r1}
- Then competition within the same location $r1$ will make p_B affect p_A
- It is still the same idea of BLP to measure potential market power from competition, but much more refined

IV: Markup Shifters or Cost Shifter

Fan-style IV but with location-specific pricing

- This IV uses rival product k's demographics in nearby market
- This IV requires us not only to observe demographics at the market level, but also at the product level

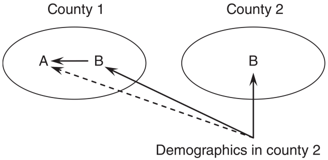


FIGURE 2. INSTRUMENTS

- Assume the newspaper now has prices that vary by year and location due to Big Data
- Consumer demographics for newspaper B in county 2 is markup shifter for product B in $r2$ affecting p_B^{r2} but mc may still be affected
- E.g., if the price of B in $r2$ increase b/c income rises in $r2$, it may create some economies of scale for the firm to serve B to both markets
- Then mc_B in location $r1$ will be affected, which in turn affect p_B^{r1}
- Then competition within the same location $r1$ will make p_B^{r1} affect p_A^{r1}
- In this case, this IV combines BLP IV, Hausman IV, and Waldfogel IV
- I would still think of it as a refined BLP IV but it has some cost-shifter ideas in it

Outline

- Introduction to Discrete Choice ✓
- Introduction to Simple Multinomial Discrete Choice with Aggregate Data ✓
- Output from Demand Estimation: Elasticities ✓
- IV ✓
- Key Practitioner Guide & Online Resource

Resource

The GOAT

- Mixtape Demand Estimation Session by Jeff Gortmaker https://github.com/Mixtape-Sessions/Demand-Estimation
- Chris Conlon's Grad IO page https://chrisconlon.github.io/gradio.html
- Train's e-book https://eml.berkeley.edu/books/choice2.html
- For all of them, simply read the part about multinomial logit is sufficient for now

Recent Literature Survey

- Berry & Haile (2021) "Foundations of Demand Estimation"
- Gandhi & Nevo (2021) "Empirical Models of Demand and Supply in Differentiated Products Industries"
- Reading them does not require an understanding of the random coefficient discrete choice of demand. You can skip the part that has random utilities, i.e., the α and β vary by individual *i*s. Focus on the big picture and IVs for now.