Lecture 3: Aggregate demand models

Jean-François Houde
UW-Madison

November 14, 2021

Introduction: Omitted attributes

- A motivating example: Trajtenberg (1989)
- **Research question:** What is the value of technology innovation?
  - *Case study:* Diffusion of CT scans
- **Demand model:**

$$\sigma_{jt} = \frac{\exp(g(x_{jt}) + \alpha(y_i - p_{jt}))}{\sum_{j' \in J_t} \exp(g(x_{jt}) + \alpha(y_i - p_{jt}))} = \frac{\exp(g(x_{jt}) - \alpha p_{jt}))}{\sum_{j' \in J_t} \exp(g(x_{j't}) - \alpha p_{j't})}$$

where z_{jt} is a vector characteristics and p_{jt} is the "residual" price of j (hedonic).

Introduction: Omitted attributes

- A motivating example: Trajtenberg (1989)
- **Research question:** What is the value of technology innovation?
  - *Case study:* Diffusion of CT scans
- **Demand model:**

$$\sigma_{jt} = \frac{\exp(g(x_{jt}) + \alpha(y_i - p_{jt}))}{\sum_{j' \in J_t} \exp(g(x_{jt}) + \alpha(y_i - p_{jt}))} = \frac{\exp(g(x_{jt}) - \alpha p_{jt}))}{\sum_{j' \in J_t} \exp(g(x_{j't}) - \alpha p_{j't})}$$

where z_{jt} is a vector characteristics and p_{jt} is the "residual" price of j (hedonic).

- Value of technological progress:

$$\begin{aligned}\Delta W_t &= \frac{1}{\hat{\alpha}} \times \left[\ln \left(\sum_{j \in J_t} \exp(\hat{g}(x_{jt}) - \hat{\alpha} p_{jt}) \right) \right. \\ &\qquad \left. - \ln \left(\sum_{j \in J_{t-1}} \exp(\hat{g}(x_{j,t-1}) - \hat{\alpha} p_{j,t-1}) \right) \right]\end{aligned}$$

Introduction: Omitted attributes

	1976	1977	1978	1979	1980	1981
RPRICE	11.252	.993	1.020	.485	.695	−.277
	(6.4)	(4.8)	(4.8)	(1.8)	(2.4)	(−2.5)
SPEED	−2.292	2.138	4.624	−8.669	11.347	−7.504
	(−7.3)	(2.8)	(1.0)	(−1.5)	(2.0)	(−.5)
SPEED^2	.236	−1.264	−8.283	31.292	−34.838	74.161
	(4.0)	(−3.4)	(−.6)	(1.9)	(−1.6)	(1.4)
RESOL	69.107	9.113	−34.126	−15.283	−18.129	32.877
	(7.3)	(2.4)	(−6.3)	(−5.0)	(−3.6)	(−3.9)
RESOL^2	−23.360	−2.533	15.096	6.291	7.738	−24.028
	(−7.6)	(−1.5)	(5.8)	(3.8)	(2.7)	(−4.2)
RTIME	−3.931	5.082	2.385	3.288	3.161	−2.591
	(−5.3)	(7.0)	(2.0)	(3.3)	(2.8)	(−2.8)
RTIME^2	1.054	−2.370	−1.511	−1.401	−2.093	5.560
	(4.5)	(−6.7)	(−2.0)	(−2.1)	(−2.2)	(3.9)
$\rho^2 = 1 - [L(\beta^*)/L(\beta^0)]$	.29	.12	.16	.16	.20	.14
$\text{Corr}(\pi^*, \pi)$	.999	.877	.900	.870	.722	.547
	(.0001)	(.0001)	(.0001)	(.0001)	(.0024)	(.082)
Number of scanners	8	15	16	16	15	11
Number of observations	285	324	164	177	193	153

NOTE.—Asymptotic t-values are in parentheses.

Introduction: Omitted attributes

	1976	1977	1978	1979	1980	1981
RPRICE	11.252	.993	1.020	.485	.695	−.277
	(6.4)	(4.8)	(4.8)	(1.8)	(2.4)	(−2.5)
SPEED	−2.292	2.138	4.624	−8.669	11.347	−7.504
	(−7.3)	(2.8)	(1.0)	(−1.5)	(2.0)	(−.5)
SPEED^2	.236	−1.264	−8.283	31.292	−34.838	74.161
	(4.0)	(−3.4)	(−.6)	(1.9)	(−1.6)	(1.4)
RESOL	69.107	9.113	−34.126	−15.283	−18.129	32.877
	(7.3)	(2.4)	(−6.3)	(−5.0)	(−3.6)	(−3.9)
RESOL^2	−23.360	−2.533	15.096	6.291	7.738	−24.028
	(−7.6)	(−1.5)	(5.8)	(3.8)	(2.7)	(−4.2)
RTIME	−3.931	5.082	2.385	3.288	3.161	−2.591
	(−5.3)	(7.0)	(2.0)	(3.3)	(2.8)	(−2.8)
RTIME^2	1.054	−2.370	−1.511	−1.401	−2.093	5.560
	(4.5)	(−6.7)	(−2.0)	(−2.1)	(−2.2)	(3.9)
$\rho^2 = 1 - [L(\beta^*)/L(\beta^0)]$	.29	.12	.16	.16	.20	.14
Corr(π^*, π)	.999	.877	.900	.870	.722	.547
	(.0001)	(.0001)	(.0001)	(.0001)	(.0024)	(.082)
Number of scanners	8	15	16	16	15	11
Number of observations	285	324	164	177	193	153

NOTE.—Asymptotic *t*-values are in parentheses.

- Upward sloping demand!
- **Omitted variable bias:** Unobserved quality of new scanners

Data: Aggregate demand and characteristics

- **Data:** Panel of market shares and product characteristics:

$$\{s_t, p_t, x_t\}_{t=1,\ldots,T}$$

  where t indexes a market, $x_t = \{x_{jt,1}, \ldots, x_{jt,K}\}_{j=1,\ldots,n_t}$ is a matrix of observed characteristics, and $\{p_t, s_t\} = \{p_{jt}, s_{jt}\}_{j=1,\ldots,n_t}$ is a matrix of endogenous prices and market shares.
- s_{jt} denotes the observed share of consumers choosing j in market t (e.g. geography, time period, etc)
- Measurement?

$$s_{jt} = \frac{Q_{jt}}{M_t}$$

  where M_t is the number of potential buyers.
- I will use w_{jt} to denote a vector of price instruments.

Incorporating Omitted Attributes: Logit case

- Relaxing the conditional independence assumption (Berry 1994):

$$V_{ijt} = x_{jt}\beta - \alpha p_{jt} + \xi_{jt} + \epsilon_{ijt}$$

  where ξ_{jt} is unobserved by the econometrician.
- This is a clear violations of the assumption and residuals are independent of (x, p):

$$E(\xi_{jt} + \epsilon_{ijt} | x_{jt}, p_{jt}) \neq 0$$

Incorporating Omitted Attributes: Logit case

- Relaxing the conditional independence assumption (Berry 1994):

$$V_{ijt} = x_{jt}\beta - \alpha p_{jt} + \xi_{jt} + \epsilon_{ijt}$$

  where ξ_{jt} is unobserved by the econometrician.

- This is a clear violations of the assumption and residuals are independent of (x, p):

$$E(\xi_{jt} + \epsilon_{ijt} | x_{jt}, p_{jt}) \neq 0$$

- MLE can identify (at most) a product/market fixed-effect:

$$\delta_{jt} = x_{jt}\beta - \alpha p_{jt} + \xi_{jt}$$

  Normalization: $\delta_{0t} = 0$.

- How?

$$\hat{\delta}_{jt} = \sigma_{jt}^{-1}(s_t) = \ln s_{jt} - \ln s_{0t} = \text{Log odds ratio}$$

Omitted attributes: Instrumental Variables

- Identification of (β, α)? Instrumental variables.

Omitted attributes: Instrumental Variables

- Identification of (β, α)? Instrumental variables.
- **Assumptions:**
  - Exogenous characteristics: $E[\xi_{jt}|x_t] = 0$
  - Price instrument (e.g. cost shifter): $E[\xi_{jt}|z_{jt}] = 0$
- Notation:
  - $y_{jt} = \ln s_{jt} - \ln s_{0t}$
  - Matrix of characteristics: Y, X and Z (X includes price)

Omitted attributes: Instrumental Variables

- Identification of (β, α)? Instrumental variables.
- **Assumptions:**
  - Exogenous characteristics: $E[\xi_{jt}|x_t] = 0$
  - Price instrument (e.g. cost shifter): $E[\xi_{jt}|z_{jt}] = 0$
- Notation:
  - $y_{jt} = \ln s_{jt} - \ln s_{0t}$
  - Matrix of characteristics: Y, X and Z (X includes price)
- Estimation: GMM

$$\text{Moment condition: } E\left[\xi_{jt} \cdot z_{jt}\right] = 0$$

$$\text{Empirical moments: } \frac{1}{n} \sum_{j,t} \underbrace{(y_{jt} - x_{jt}\beta + \alpha p_{jt}) \cdot z_{jt}}_{g_{jt}(\theta)} = 0$$

where $n = \sum_t J_t$ is the number of observations and $\theta = (\alpha, \beta)$

- Empirical moments in matrix form:

$$\bar{g}_n(\theta) = n^{-1} Z'(Y - X\theta)$$

Omitted attributes with Non-IIA Demand: Nested-Logit

Source: Berry (1994)

- Lets drop p_{jt} (Exogenous characteristics)

Omitted attributes with Non-IIA Demand: Nested-Logit

Source: Berry (1994)

- Lets drop p_{jt} (Exogenous characteristics)
- **Example:**
  - G segments

$$V_{ijt} = \delta_{jt} + \nu_{ig} + (1-\lambda)\epsilon_{ijt} \text{ If } j \in g$$

λ is nested-logit parameter (i.e. correlation in errors).

  - Demand:

$$\sigma_j(\delta_t, G_t) = \frac{\exp(\delta_{jt}/(1-\lambda))}{H_g\left[\sum_{g'} H_{g'}\right]}$$

where $H_g = \ln\left[\sum_{j\in g}\exp(\delta_{jt}/(1-\lambda))\right]$ is the inclusive-value of segment g.

Omitted attributes with Non-IIA Demand: Nested-Logit

Source: Berry (1994)

- Lets drop p_{jt} (Exogenous characteristics)
- **Example:**
  - G segments

$$V_{ijt} = \delta_{jt} + \nu_{ig} + (1-\lambda)\epsilon_{ijt} \text{ If } j \in g$$

    λ is nested-logit parameter (i.e. correlation in errors).
  - Demand:

$$\sigma_j(\delta_t, G_t) = \frac{\exp(\delta_{jt}/(1-\lambda))}{H_g\left[\sum_{g'} H_{g'}\right]}$$

    where $H_g = \ln\left[\sum_{j\in g} \exp(\delta_{jt}/(1-\lambda))\right]$ is the inclusive-value of segment g.
- Inverse-demand:

$$s_{jt} = \sigma_j(\delta_t, G_t; \lambda) \rightarrow \delta_{jt} = \ln s_{jt} - \ln s_{0t} - \lambda s_{j|g,t} = \sigma_j^{-1}(s_t, G_t; \lambda)$$

where $s_{j|g,t}$ is the conditional share of j in segment g.

Nested-Logit

Source: Berry (1994)

- **Implication:** The inverse demand takes the form of a linear regression

$$y_{jt} = \ln s_{jt}/s_{0t} = x_{jt}\beta + \lambda s_{j|g,t} + \xi_{jt}$$

- **Implication:** $(\beta^{ols}, \lambda^{ols})$ is biased. Why?
  - The popularity of j in segment g is function of ξ_{jt}
  - $\hat{\lambda}^{ols}$ is biased upward: Too much correlation in taste
  - This biased is present wether or not characteristics are exogenous

Nested-Logit

Source: Berry (1994)

- **Implication:** The inverse demand takes the form of a linear regression

$$y_{jt} = \ln s_{jt}/s_{0t} = x_{jt}\beta + \lambda s_{j|g,t} + \xi_{jt}$$

- **Implication:** $(\beta^{ols}, \lambda^{ols})$ is biased. Why?
  - The popularity of j in segment g is function of ξ_{jt}
  - $\hat{\lambda}^{ols}$ is biased upward: Too much correlation in taste
  - This biased is present wether or not characteristics are exogenous
- **Omitted variable bias:** When a product is popular in a segment. Two possible reasons:
  - **Large λ:** Rival products have unfavorable attributes (or fewer options)
  - **Zero λ:** Product j has large *unobserved* quality
- Without further assumptions on (G_t, x_t) it impossible consistently identify λ

Nested-Logit

Source: Berry (1994)

- **Assumption:** The residual quality of j is independent of the menu of characteristics $x_t = \{x_{1t}, \ldots, x_{J_t,t}\}$ available in market t

$$E\left[\xi_{jt} | x_t\right] = 0$$

Nested-Logit

Source: Berry (1994)

- **Assumption:** The residual quality of j is independent of the menu of characteristics $x_t = \{x_{1t}, \ldots, x_{J_t,t}\}$ available in market t

$$E\left[\xi_{jt} | x_t\right] = 0$$

- This conditional moment restriction (CMR) was introduced by Berry (1994) and Berry et al. (1995)

Nested-Logit

Source: Berry (1994)

- **Assumption:** The residual quality of j is independent of the menu of characteristics $x_t = \{x_{1t}, \ldots, x_{J_t,t}\}$ available in market t

$$E\left[\xi_{jt} | x_t\right] = 0$$

- This conditional moment restriction (CMR) was introduced by Berry (1994) and Berry et al. (1995)
- This is not as strong as it seems like...
- The quality index δ_{jt} is linear in ξ_{jt} and so we can condition of rich fixed-effects. E.g.:

$$\delta_{jt} = x_{jt}\beta + \mu_j + \tau_t + \xi_{jt} \rightarrow E\left[\xi_{jt} | x_t, \mu_j, \tau_t\right] = 0$$

Nested-Logit

Source: Berry (1994)

- **Assumption:** The residual quality of j is independent of the menu of characteristics $x_t = \{x_{1t}, \dots, x_{J_t,t}\}$ available in market t

$$E\left[\xi_{jt} | x_t\right] = 0$$

- This conditional moment restriction (CMR) was introduced by Berry (1994) and Berry et al. (1995)
- This is not as strong as it seems like...
- The quality index δ_{jt} is linear in ξ_{jt} and so we can condition of rich fixed-effects. E.g.:

$$\delta_{jt} = x_{jt}\beta + \mu_j + \tau_t + \xi_{jt} \rightarrow E\left[\xi_{jt} | x_t, \mu_j, \tau_t\right] = 0$$

- **Example:** Moment conditions with option fixed-effects:

$$E\left[\xi_{jt} \cdot \dot{z}_{jt}\right] = E\left[\left(\dot{y}_{jt} - \dot{x}_{jt}\beta - \lambda \dot{\ln s_{j|g,t}}\right) \cdot \dot{z}_{jt}\right]$$

where $\dot{x}_{jt} = x_{jt} - \bar{x}_j$ is the "within-transformation" of characteristics.

Nested-Logit

Source: Berry (1994)

- Examples of relevant IVs for λ:
  - Number of products in the same nest
  - Characteristics of products in the same nest

Nested-Logit

Source: Berry (1994)

- Examples of relevant IVs for λ:
  - Number of products in the same nest
  - Characteristics of products in the same nest
- Adding back prices:

$$y_{jt} = \ln s_{jt}/s_{0t} = x_{jt}\beta - \alpha p_{jt} + \lambda s_{j|g,t} + \xi_{jt}$$

- **Takeaway:** We need two independent sources of exogenous variation
  - IV for p: Cost/markup shifter
  - IV for s: Characteristics of rival products

General Case: Mixed-logit demand system

- Indirect utility function (Nevo 2001):

$$V_{ijt} = \begin{cases} \sum_{k=1}^{K} \beta_{i,k} x_{jt,k} - \alpha_i p_{jt} + \xi_{jt} + \epsilon_{ij} & \text{If } j \neq 0 \\ \epsilon_{i0} & \text{Else.} \end{cases}$$

$$\beta_{i,k} = \beta_k + z_i \pi_k + \nu_{i,k}$$

$$\alpha_i = \alpha + z_i \pi_p + \nu_{i,p}$$

$$z_i \sim D_t(\cdot) \text{ (known) and } \nu_i \sim F(\nu_i; \lambda) \text{ (unknown)}$$

Average utility: $\delta_{jt} = x_{jt}\beta - \alpha p_{jt} + \xi_{jt}$

General Case: Mixed-logit demand system

- Indirect utility function (Nevo 2001):

$$V_{ijt} = \begin{cases} \sum_{k=1}^{K} \beta_{i,k} x_{jt,k} - \alpha_i p_{jt} + \xi_{jt} + \epsilon_{ij} & \text{If } j \neq 0 \\ \epsilon_{i0} & \text{Else.} \end{cases}$$

$$\beta_{i,k} = \beta_k + z_i \pi_k + \nu_{i,k}$$

$$\alpha_i = \alpha + z_i \pi_p + \nu_{i,p}$$

$$z_i \sim D_t(\cdot) \text{ (known) and } \nu_i \sim F(\nu_i; \lambda) \text{ (unknown)}$$

$$\text{Average utility: } \delta_{jt} = x_{jt}\beta - \alpha p_{jt} + \xi_{jt}$$

- Characteristics vector: $x_{jt} = \left\{ x_{jt}^{(1)}, x_{jt}^{(2)} \right\}$.
  - Attributes with common valuation: $x_{jt}^{(1)}$
  - Attributes with heterogenous valuation: $x_{jt}^{(2)}$
- Parameters:
  - *Linear:* Mean utility parameters (β, α)
  - *Non-linear:* Demographic weights (π) and random-coefficients (λ)

Baseline Model: Exogenous Characteristics

- Consider first a model without prices and without demographics characteristics

Baseline Model: Exogenous Characteristics

- Consider first a model without prices and without demographics characteristics
- **Demand:** Linear random-coefficient with T1EV random utility shocks

$$\sigma_j\left(\delta_t, x_t^{(2)}; \lambda\right) = \int \frac{\exp\left(\delta_{jt} + \nu_i^T x_{jt}^{(2)}\right)}{1 + \sum_{j'=1}^{J_t} \exp\left(\delta_{j't} + \nu_i^T x_{j't}^{(2)}\right)} dF(\nu_i|\lambda)$$

where $\delta_{jt} = \beta_0 + x_{jt}^{(1)}\beta_1 + x_{jt}^{(2)}\beta_2 + \xi_{jt}$.

Baseline Model: Exogenous Characteristics

- Consider first a model without prices and without demographics characteristics
- **Demand:** Linear random-coefficient with T1EV random utility shocks

$$\sigma_j\left(\delta_t, x_t^{(2)};\lambda\right) = \int \frac{\exp\left(\delta_{jt} + \nu_i^T x_{jt}^{(2)}\right)}{1 + \sum_{j'=1}^{J_t} \exp\left(\delta_{j't} + \nu_i^T x_{j't}^{(2)}\right)} dF(\nu_i|\lambda)$$

where $\delta_{jt} = \beta_0 + x_{jt}^{(1)}\beta_1 + x_{jt}^{(2)}\beta_2 + \xi_{jt}$.
- The **residual** of the model is obtained from the inverse-demand function:

$$\rho_j(s_t, x_t;\theta) = \sigma_j^{-1}\left(s_t, x_t^{(2)};\lambda\right) - x_{jt}\beta, \quad \text{where } \theta = (\beta,\lambda).$$

Baseline Model: Exogenous Characteristics

- Consider first a model without prices and without demographics characteristics
- **Demand:** Linear random-coefficient with T1EV random utility shocks

$$\sigma_j\left(\delta_t, x_t^{(2)}; \lambda\right) = \int \frac{\exp\left(\delta_{jt} + \nu_i^T x_{jt}^{(2)}\right)}{1 + \sum_{j'=1}^{J_t} \exp\left(\delta_{j't} + \nu_i^T x_{j't}^{(2)}\right)} dF(\nu_i|\lambda)$$

  where $\delta_{jt} = \beta_0 + x_{jt}^{(1)}\beta_1 + x_{jt}^{(2)}\beta_2 + \xi_{jt}$.
- The **residual** of the model is obtained from the inverse-demand function:

$$\rho_j(s_t, x_t; \theta) = \sigma_j^{-1}\left(s_t, x_t^{(2)}; \lambda\right) - x_{jt}\beta, \quad \text{where } \theta = (\beta, \lambda).$$

- **Important:** The model residual is a non-linear function of λ. This leads to a non-linear IV estimator

Identifying Assumption

- **Assumption:** The unobserved attribute of each product is independent of the **menu**, x_t, of characteristics available in market t,

$$E[\xi_{jt}|x_t] = 0 \quad \text{(CMR)}.$$

- In practice, the model is estimated using a finite number (L) of unconditional moment restrictions, $A_j(x_t)$ (aka instruments):

$$\begin{aligned} & E\left[\rho_j(s_t, x_t; \theta^0) \cdot A_j(x_t)\right] = 0 \\ \leftrightarrow \quad & E\left[\left(\sigma_j^{-1}\left(s_t, x_t^{(2)}; \lambda^0\right) - x_{jt}\beta\right) \cdot A_j(x_t)\right] = 0. \end{aligned}$$

- **Questions:**
  - How to choose $A_j(x_t)$?
  - How to estimate (β, λ)?

How should we choose the instruments?

- True model: Random-coefficient with exogenous attributes

$$\sigma_j(\delta_t, x_t^{(2)}|\lambda) = \int \frac{\exp(\delta_{jt} + \lambda\eta_i x_{jt}^{(2)})}{1 + \sum_{j'} \exp(\delta_{j't} + \lambda\eta_i x_{j't}^{(2)})} dF(\eta_i|\lambda)$$
$$\Rightarrow \quad x_{jt}\beta + \xi_{jt} \equiv \delta_{jt} = \sigma_j^{-1}(s_t, x_t^{(2)}|\lambda)$$

- Wrong model: Logit

$$\sigma_j^{-1}\left(s_t, x_t^{(2)}|\lambda = 0\right) = \ln s_{jt}/s_{0t} = x_{jt}\beta + \underbrace{\Delta_{jt} + \xi_{jt}}_{r_{jt} = \text{Logit residual}}$$

where $\Delta_{jt} = \sigma_j^{-1}(s_t, x_t^{(2)}|\lambda = 0) - \sigma_j^{-1}(s_t, x_t^{(2)}|\lambda)$ (omitted variable)

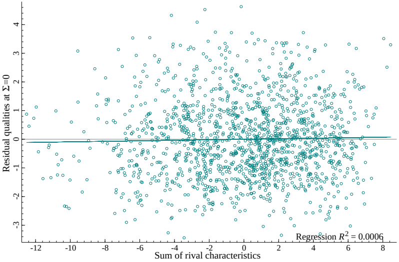
- Identification: $A_j(x_t)$ is a strong and valid instrument if

$$E[\xi_{jt}A_j(x_t)] = 0 \text{ and } E[r_{jt}A_j(x_t)] \neq 0$$

Identification in a Picture

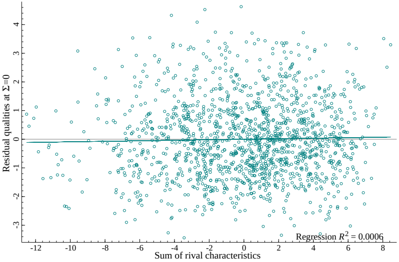
Identification in a Picture

(A) IV: Sum of rivals' characteristics

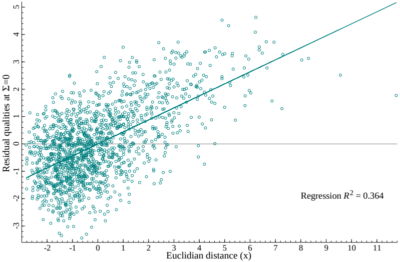


Identification in a Picture

(A) IV: Sum of rivals' characteristics



(B) IV: Euclidean distance in x



- **Interpretation:** Products that have few close substitutes (↑ distance) have large market-share (at the true λ)
- Under logit ($\lambda = 0$), isolated products are predicted to have higher quality (i.e. $\uparrow \Delta_{jt}$)
- Clear violation of the moment condition

Instrument selection: General result

- An instrument function is "strong" if it is a good predictor of the inverse-demand function:

$$E[\sigma^{-1}(s_t, x^{(2)}|\lambda)|x_t] \approx A_j(x_t)\gamma$$

This corresponds to the reduced-form of the model.

Instrument selection: General result

- An instrument function is "strong" if it is a good predictor of the inverse-demand function:

$$E[\sigma^{-1}(s_t, x^{(2)}|\lambda)|x_t] \approx A_j(x_t)\gamma$$

  This corresponds to the reduced-form of the model.

- **Proposition (Gandhi and Houde 2020):** If the distribution of $\{\xi_j\}_{j=1,\dots,n}$ is exchangeable (conditional on x_{jt}), then the reduced form can be written as

$$E\left[\sigma_j^{-1}\left(\boldsymbol{s}, \boldsymbol{x}^{(2)}; \boldsymbol{\lambda}^0\right)|\boldsymbol{x}\right] = g\left(\boldsymbol{d}_j\right)$$

  where g is a **symmetric** function of the state vector.

Instrument selection: General result

- An instrument function is "strong" if it is a good predictor of the inverse-demand function:

$$E[\sigma^{-1}(s_t, x^{(2)}|\lambda)|x_t] \approx A_j(x_t)\gamma$$

This corresponds to the reduced-form of the model.

- **Proposition (Gandhi and Houde 2020):** If the distribution of $\{\xi_j\}_{j=1,\dots,n}$ is exchangeable (conditional on x_{jt}), then the reduced form can be written as

$$E\left[\sigma_j^{-1}\left(\boldsymbol{s}, \boldsymbol{x}^{(2)}; \boldsymbol{\lambda}^0\right)|\boldsymbol{x}\right] = g\left(\boldsymbol{d}_j\right)$$

where g is a **symmetric** function of the state vector.

- **Implication:** The optimal instrument is a function of the distribution of characteristics differences (aka Differentiation IVs)

Why is symmetry of the reduced-form useful?

- The reduced form is a **symmetric** function of characteristic differences:
$$E\left[\sigma_j^{-1}\left(\boldsymbol{s}, \boldsymbol{x}^{(2)}; \boldsymbol{\lambda}^0\right)|\boldsymbol{x}\right] = g\left(\boldsymbol{d}_j\right)$$
- **Definition:** $g(d_1, d_2)$ is a symmetric function if $g(d_1, d_2) = g(d_2, d_1)$.

Why is symmetry of the reduced-form useful?

- The reduced form is a **symmetric** function of characteristic differences:
$$E\left[\sigma_j^{-1}\left(\boldsymbol{s}, \boldsymbol{x}^{(2)}; \boldsymbol{\lambda}^0\right)|\boldsymbol{x}\right] = g\left(\boldsymbol{d}_j\right)$$
- **Definition:** $g(d_1, d_2)$ is a symmetric function if $g(d_1, d_2) = g(d_2, d_1)$.
- **Example:** Single dimension $d_{jt} = \{x_{1t} - x_{jt}, x_{2t} - x_{jt}, \ldots, x_{J_t,t} - x_{jt}\}$
  - Second-order approximation of $g(d)$:

$$\begin{aligned} g(d_{jt}) &\approx \sum_{j'} \gamma^1_{j'} d_{jt,j'} + \sum_{j'} \gamma^2_{j'} (d_{jt,j'})^2 + \gamma^3 \left(\sum_{j'} d_{jt,j'}\right)^2 \\ &= \gamma^1 \left(\sum_{j'} d_{jt,j'}\right) + \gamma^2 \left(\sum_{j'} (d_{jt,j'})^2\right) + \gamma^3 \left(\sum_{j'} d_{jt,j'}\right)^2 \end{aligned}$$

Implication 1: Polynomial Basis

- Single dimension measures of differentiation

$$\text{Quadratic: } A_j(x_t) = \sum_{j'} \left(d^k_{jt,j'} \right)^2$$

Note: $\sqrt{z_{jt,k}}$ is the Euclidian distance between product j and its rivals in market t along dimension k.

- Adding interaction terms:

$$\text{Covariance: } A_j(x_t) = \sum_{j'} d^k_{jt,j'} \times d^l_{jt,j'}$$

Implication 2: Histogram Basis

- **Note:** This approach is advisable only in very large samples (+large choice-sets), and when the goal is to estimate a very flexible distribution of RCs.
- Single dimension measure of differentiation = Number of rivals in discrete bins

$$A_j(x_t) = \left\{ \sum_{j'} 1\left(d^k_{jt,j'} < \kappa_l\right) \right\}_{l=1,\dots,L}$$

- Multi-dimension measure of differentiation:

$$A_j(x_t) = \left\{ \sum_{j'} 1\left(d^k_{jt,j'} < \kappa_l\right) 1\left(d^{k'}_{jt,j'} < \kappa_{l'}\right) \right\}_{l=1,\dots,L,l'=1,\dots,L}$$

Implication 3: Local Basis

- **Note:** In most parametric models, the inverse demand is function of characteristics of **close-by** rivals. Therefore, in the previous histogram basis, we should be focussing on "local" rivals.
- Single dimension measure of differentiation = Number of nearby rivals along each dimension

$$A_j(x_t) = \sum_{j'} 1\left(|d^k_{jt,j'}| < \kappa_k\right), \text{ e.g. } \kappa_k = sd(x_{jt,k})$$

- Multi-dimension measure of differentiation:

$$A_j(x_t) = \sum_{j'} 1\left(|d^k_{jt,j'}| < \kappa_k\right) \times d_{jt,l}, \text{ e.g. } \kappa_k = sd(x_{jt,k})$$

- When $x_{jt,k}$ is discrete, this basis function boils down to the familiar Nested-logit IVs.
    - Number of competitors and characteristics of rivals within segment

Differentiation IVs with Endogenous Prices

Example with cost shifter

1. First-stage price regression:

$$\hat{p}_{jt} = \hat{\pi}_0 + \hat{\pi}_1 x_{jt} + \hat{\pi}_2 \omega_{jt}$$

   This could be richer/more non-linear. Goal: Provide a good fit using (x, w)

Differentiation IVs with Endogenous Prices

Example with cost shifter

1. First-stage price regression:

$$\hat{p}_{jt} = \hat{\pi}_0 + \hat{\pi}_1 x_{jt} + \hat{\pi}_2 \omega_{jt}$$

   This could be richer/more non-linear. Goal: Provide a good fit using (x, w)

2. Differentiation IV: Quadratic

$$\sum_{j'} \left(d^{\hat{p}}_{jt,j'} \right)^2 \text{ and } \sum_{j'} \left(d^{\hat{p}}_{jt,j'} \right)^2 \cdot \boldsymbol{d}_{jt,j'}$$

   where $\boldsymbol{d}_{jt,j'} = (d^x_{jt,j'}, d^{\hat{p}}_{jt,j'})$.

Differentiation IVs with Endogenous Prices

Example with cost shifter

1. First-stage price regression:

$$\hat{p}_{jt} = \hat{\pi}_0 + \hat{\pi}_1 x_{jt} + \hat{\pi}_2 \omega_{jt}$$

   This could be richer/more non-linear. Goal: Provide a good fit using (x, w)

2. Differentiation IV: Quadratic

$$\sum_{j'} \left(d^{\hat{p}}_{jt,j'}\right)^2 \text{ and } \sum_{j'} \left(d^{\hat{p}}_{jt,j'}\right)^2 \cdot \boldsymbol{d}_{jt,j'}$$

   where $\boldsymbol{d}_{jt,j'} = (d^x_{jt,j'}, d^{\hat{p}}_{jt,j'})$.

3. Differentiation IV: Local

$$\sum_{j'} \left(|d^{\hat{p}}_{jt,j'}| < \text{sd}(\hat{p}_{jt})\right) \text{ and } \sum_{j'} \left(|d^{\hat{p}}_{jt,j'}| < \text{sd}(\hat{p}_{jt})\right) \cdot \boldsymbol{d}_{jt,j'}$$

How can we estimate the model?

- Non-linear GMM estimator:
  - Moment conditions:

$$\bar{g}(\beta,\theta) = \frac{1}{n}\sum_{j,t} Z_{jt} \underbrace{\left(\sigma_j^{-1}(s_t, x_t^{(2)}, p_t|\lambda) - X_{jt}\beta\right)}_{\text{Residual: } \rho_{jt}(\beta,\lambda)} = 0$$

  - Nested fixed-point optimization problem:

$$\begin{aligned} \min_{\lambda} \quad & \rho(\lambda)^T Z W^{-1} Z^T \rho(\lambda) \\ s.t. \quad & \rho_{jt}(\lambda) = \sigma_j^{-1}(s_t, x_t^{(2)}, p_t|\lambda) - X_{jt}\beta(\lambda) \\ & \beta(\lambda) = \text{Linear IV estimate given } \lambda \end{aligned}$$

  - Note: Expressing β as a function of λ greatly simplifies the optimization problem (i.e. $dim(\lambda) << dim(\beta)$).

Three computational challenges

- Numerical optimization problem: BFGS or Simplex (relatively simple when IVs are strong)
- Numerical integration:

$$\begin{aligned}
\sigma_j(\delta_t, x_t^{(2)}, p_t|\lambda) &= \int \frac{\exp(\delta_{jt} + \alpha_i p_{jt} + x^{(2)}\nu_i)}{1 + \sum_{j'} \exp(\delta_{j',t} + \alpha_i p_{j',t} + x_{j',t}^{(2)}\nu_i)} dF(\alpha_i, \nu_i|\lambda) \\
&\approx \sum_r w_r \frac{\exp(\delta_{jt} + \alpha_r p_{jt} + x^{(2)}\nu_r)}{1 + \sum_{j'} \exp(\delta_{j',t} + \alpha_r p_{j',t} + x_{j',t}^{(2)}\nu_r)}
\end{aligned}$$

  Example: Monte-carlo simulation ($w_r = 1/R$), quadrature method ($w_r =$ gaussian weights).

- Non-linear equation solution (fixed-point):

$$\delta_{jt} = \sigma_j^{-1}(s_t, x_t^{(2)}, p_t|\lambda)$$

Pseudo-code

```
/* GMM objective function */
gmm_obj(const vP, const adFunc, const avScore, const amHessian)
{
  /* Invert demand */
  inverse(&vDelta0,vP);
  /* Quality decomposition */
  decl vWithinDelta0=within(vDelta0,vFEid);
  decl vLParam=ivreg(vWithinDelta0,mX,mIV,A);
  decl vXi=vWithinDelta0-mX*vLParam;

  /* GMM objective function */
  mG=vXi.*mIV;
  decl scale=100;
  decl g=sumc(mG);
  if(isnan(vDelta0)==1) adFunc[0]=.NaN;
  else adFunc[0]=double(-g*A*g'/scale);
  if(avScore) {
    /* Score */
    decl mJacobian;
    jacobian(&mJacobian,vDelta0,vP);
    mJacobian=within(mJacobian,vFEid);
    decl dG=(mMx*mJacobian)'mIV;
    decl vScore=-2*dG*A*g'/scale;
    avScore[0]=vScore;
  }
  return 1;
}
```

Demand inversion: Fixed-point algorithm

- **Algorithm 1 (Berry et al. (1995)):** The following function is a contraction mapping

$$\delta_{jt}^{i} = \Gamma(\delta_t^{i-1}) \equiv \delta_{jt}^{-1} + \underbrace{\left(\ln s_{jt} - \ln \sigma_j(\delta_t^{i-1})\right)}_{\text{step size}}$$

1. Initial guess: δ_{jt}^{0}
2. Iteration i:
$$\delta_{jt}^{i} = \delta_{jt}^{i-1} + \left(\ln s_{jt} - \ln \sigma_j(\delta_t^{i-1})\right)$$
3. Stop if $||\ln s_{jt} - \ln \sigma_j(\delta_t^{i-1})|| < \varepsilon$. Else repeat step 2.

Demand inversion: Fixed-point algorithm

- **Algorithm 2:** Newton's nonlinear root-finding algorithm
  - Univariate example:

$$f(x) = 0$$

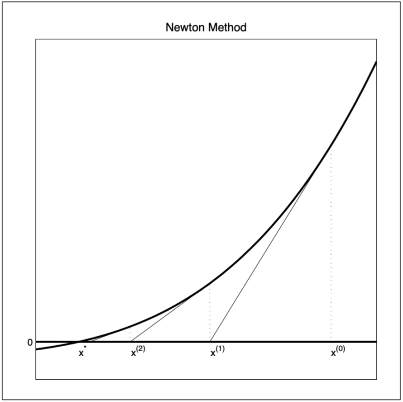
  - Initial guess: x^0
  - Linear approximation:

$$0 = f(x) \approx f(x^0) + f'(x^0)(x^1 - x^0)$$

$$x^1 = x^0 - f'(x^0)f(x^0)$$

  - Iteration k:

$$x^k = x^k - f'(x^k)f(x^k)$$



Demand inversion: Fixed-point algorithm

- **Algorithm 2:** Newton's nonlinear root-finding algorithm
  - "Zero" functions ($J_t \times 1$):

$$f(\delta_t) = \ln s_t - \ln \sigma(\delta_t)$$

  - Jacobian matrix:

$$Df(\delta) = -\left[\frac{\partial \sigma_j(\delta_t)}{\partial \delta_{kt}}\right] / \sigma(\delta_t)$$

where

$$\frac{\partial \sigma_j(\delta_t)}{\partial \delta_{kt}} = \begin{cases} \sum_r w_r \sigma_j(\delta_t|\nu_r)(1-\sigma_j(\delta_t|\nu_r)) & \text{If } j = k \\ -\sum_r w_r \sigma_j(\delta_t|\nu_r)\sigma_k(\delta_t|\nu_r) & \text{If } j \neq k \end{cases}$$

  - Pseudo-code:
    1. Initial step: $\delta_t^0 = \ln s_t - \ln s_{0,t}$ (Logit)
    2. Iteration k: Demand and Jacobian calculation
$$\sigma_{jt}(\delta_t^k) \text{ and } Df(\delta^k)$$
    3. Updating step:
$$\delta^{k+1} = \delta^k - Df(\delta)^{-1} f(\delta^k)$$
    4. Stopping rule: $f(\delta^k) < \varepsilon$.

Sample code: Demand and Jacobian calculation

```
value(const aMu,const vParam,const t)
{
  decl i;
  decl rowid=aProductID[t];
  decl mMu=new matrix[rows(rowid)][rows(aDemoID[t])];
  for(i=0;i<columns(mZ);i++) mMu+=vParam[i]*(aZ[t])[i];
  aMu[0]=exp(mMu);
  return 1;
}
demand(const mMu,const aShare,const aJac,const vDelta,const t,const vParam)
{
  decl i;
  decl rowid=aProductID[t];
  decl eV=exp(vDelta[rowid]).*mMu;
  decl mS=eV./(1+sumc(eV));
  decl vShat=meanr(mS);
  if(aJac[0]) {
    decl mD=diag(meanr(mS.*(1-mS)))-setdiagonal(mS*mS'/rows(aDemoID[t]),0);
    aJac[0]=mD;
    }
  aShare[0]=vShat;
  return 1;
}
```

Sample code: Inversion algorithm (parallel)

```
parallel for(t=0;t<T;t++)
  {
    value(&mMu,vParam,t);
    rowid=aProductID[t];
    vIT[t]=0;
    f=1;
    do{
      if(norm(f)>eps1) {
        mJacobian=0;
        demand(mMu,&vShat,&mJacobian,vDelta,t,vParam);
      }
      else {
        mJacobian=1;
        demand(mMu,&vShat,&mJacobian,vDelta,t,vParam);
      }
      f=log(vShare[rowid])-log(vShat);
      if(mJacobian==0) vDelta[rowid]=vDelta[rowid]+f;
      else vDelta[rowid]=vDelta[rowid]+invert(mJacobian./vShat)*f;
      vIT[t]+=1;
    }while(norm(f)>eps && vIT[t]<maxit);
    if(norm(f)>eps) vDelta[rowid]=constant(.NaN,rowid);
  }
```

Berry, S. (1994).
Estimating discrete choice models of product differentiation.
Rand Journal of Economics 25, 242–262.

Berry, S., J. Levinsohn, and A. Pakes (1995).
Automobile prices in market equilibrium.
Econometrica: Journal of the Econometric Society 63(4), 841–890.

Gandhi, A. and J.-F. Houde (2020, January).
Measuring substitution patterns in differentiated products industries.
NBER working paper.

Miravete, E., K. Seim, and J. Thurk (2017, September).
Market power and the laffer curve.
working paper, UT Austin.

Nevo, A. (2001).
Measuring market power in the ready-to-eat cereal industry.
Econometrica 69(2), 307.

Trajtenberg, M. (1989, Apr.).
The welfare analysis of product innovations, with an application to computed tomography scanners.
The Journal of Political Economy 97(2), 444–479.