# An Empirical Model of Quantity Discounts with Large Choice Sets

Alessandro Iaria (Bristol) Ao Wang (Warwick)

EEA-ESEM, Barcelona
August 2023

# Introduction

- **Quantity discounts** are a common form of nonlinear pricing.

- Lower unit-price for larger quantities.

# Introduction

- **Quantity discounts** are a common form of nonlinear pricing.
- Lower unit-price for larger quantities.
- Enable firms to **screen between high- and low-demand consumers** ...
- ... but may be **socially inefficient** (Varian, 1992, textbook).

# Introduction

- **Quantity discounts** are a common form of nonlinear pricing.
- Lower unit-price for larger quantities.
- Enable firms to **screen between high- and low-demand consumers** ...
- ... but may be **socially inefficient** (Varian, 1992, textbook).
- Prominent in packaged goods (food, drinks, toiletry, etc.) and services (energy, telecom, public transport, etc.).
- Despite widespread diffusion and a vast theoretical literature (Anderson and Renault, 2011; Armstrong, 2016), relatively **few empirical studies**.

# Difficulties in Empirical Demand for Bundles / Multiple Units

- Practical complexity in estimation of **demand for bundles**.
    - Even with a few products, **number of bundles quickly grows large**.

# Difficulties in Empirical Demand for Bundles / Multiple Units

- Practical complexity in estimation of **demand for bundles**.
  - Even with a few products, **number of bundles quickly grows large**.
- Non-parametric estimation typically for small choice sets (Compiani, 2019).
- (Even) Parametric **estimation can be hard with many bundles**.
  - Berry et al. (2014); Iaria and Wang (2019).

# This Paper: An Empirical Model of Demand for Bundles

- By construction, each product can be part of many bundles.

$\Longrightarrow$ **Correlated preferences across bundles** (Gentzkow, 2007).

# This Paper: An Empirical Model of Demand for Bundles

- By construction, each product can be part of many bundles.

$\Longrightarrow$ **Correlated preferences across bundles** (Gentzkow, 2007).

  - Multinomial logit and nested logit cannot capture this.
  - Mixed logit can but may be hard to estimate with large choice sets.

# This Paper: An Empirical Model of Demand for Bundles

- By construction, each product can be part of many bundles.

$\Longrightarrow$ **Correlated preferences across bundles** (Gentzkow, 2007).

  - Multinomial logit and nested logit cannot capture this.
  - Mixed logit can but may be hard to estimate with large choice sets.

- **Product-Overlap Nested Logit (PONL)**: bundles belong to multiple nests.
  - More flexible than nested logit but simpler than mixed logit.

- **Concentrated 2SLS**: extension of 2SLS by Berry (1994) to PONL.
  - Convenient with large choice sets.
  - Controls for price endogeity.

# This Paper: Quantity Discounts in Carbonated Soft Drinks (CSDs)

- As part of an anti-obesity strategy, **proposal in the UK to ban from supermarkets quantity discounts** on high-salt/sugar products.

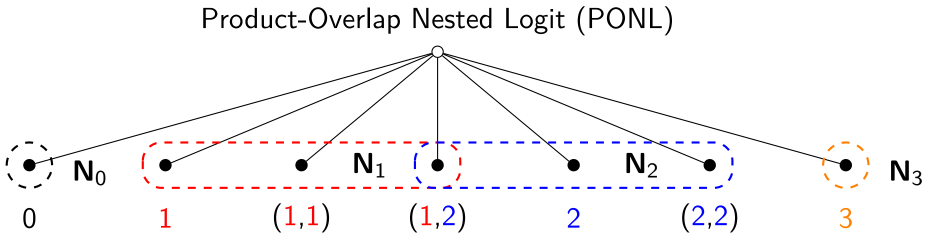
# This Paper: Quantity Discounts in Carbonated Soft Drinks (CSDs)

- As part of an anti-obesity strategy, **proposal in the UK to ban from supermarkets quantity discounts** on high-salt/sugar products.

- Use IRI data (USA, 2008-11) to estimate demand for bundles of CSDs.

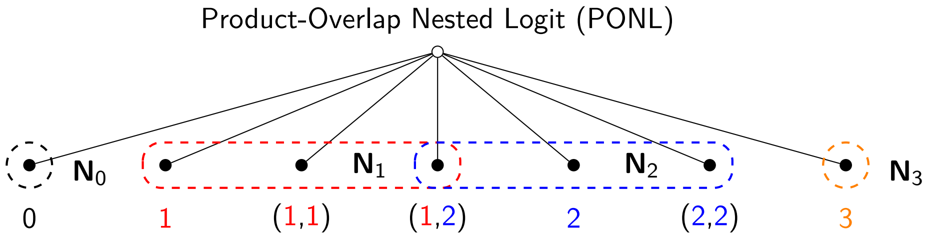
- **Simulate counterfactual linear pricing**, a constant unit-price for each product.
  - Quantity $\downarrow$ **20.7%**, profit $\downarrow$ **19.7%**, and consumer surplus $\downarrow$ **2.8%**.
  - **Added sugar intake $\downarrow$ 22.1%**.

# Product-Overlap Nested Logit (PONL): Basic Idea




- Choice set defined over bundles, and bundle **b** a collection of units of products.
- Any **b** that includes at least one unit of product $j$ belongs to nest $\mathbf{N}_j$, $j = 1, ..., J$.

# Product-Overlap Nested Logit (PONL): Basic Idea




- Choice set defined over bundles, and bundle **b** a collection of units of products.
- Any **b** that includes at least one unit of product $j$ belongs to nest $\mathbf{N}_j$, $j = 1, ..., J$.
- **Product-overlap** among bundles **determines the nests they share**.
  - $\uparrow$ similarity in product composition $\Longrightarrow \uparrow$ correlation in unobserved preferences.

# PONL Model: Average Indirect Utility and Demand Synergies

Average utility of **one unit of product** $j$ in market $t$ (Berry, 1994):

$$\delta_{tj} = \delta_j + x_{tj}\beta - \alpha p_{tj} + \xi_{tj} \tag{1}$$

while of **bundle b** in market $t$ (Gentzkow, 2007):

$$\delta_{t\mathbf{b}} = \sum_{j \in \mathbf{b}} \delta_{tj} + \Gamma_{t\mathbf{b}}. \tag{2}$$

# PONL Model: Average Indirect Utility and Demand Synergies

Average utility of **one unit of product $j$** in market $t$ (Berry, 1994):

$$\delta_{tj} = \delta_j + x_{tj}\beta - \alpha p_{tj} + \xi_{tj} \tag{1}$$

while of **bundle b** in market $t$ (Gentzkow, 2007):

$$\delta_{t\mathbf{b}} = \sum_{j \in \mathbf{b}} \delta_{tj} + \Gamma_{t\mathbf{b}}. \tag{2}$$

- $\Gamma_{t\mathbf{b}}$ is the **demand synergy**: extra-(dis)utility of joint purchase.
    - Complementarity (Gentzkow, 2007), shopping costs (Thomassen et al., 2017), preference for variety (Dubé, 2004), etc.
    - **Quantity Discounts**: $\Gamma_{t\mathbf{b}} = -\alpha \left( p_{t\mathbf{b}} - \sum_{j \in \mathbf{b}} p_{tj} \right) > 0$.

# PONL Model: Average Indirect Utility and Demand Synergies

Average utility of **one unit of product** $j$ in market $t$ (Berry, 1994):

$$\delta_{tj} = \delta_j + x_{tj}\beta - \alpha p_{tj} + \xi_{tj} \tag{1}$$

while of **bundle b** in market $t$ (Gentzkow, 2007):

$$\delta_{t\mathbf{b}} = \sum_{j \in \mathbf{b}} \delta_{tj} + \Gamma_{t\mathbf{b}}. \tag{2}$$

- $\Gamma_{t\mathbf{b}}$ is the **demand synergy**: extra-(dis)utility of joint purchase.

- Remain agnostic about $\Gamma_{t\mathbf{b}}$, a parameter to be identified and estimated.

# PONL Model: Overlapping Nests and Berry (1994)

The PONL inverse demand for **bundle b** in market $t$

$$\ln(\mathcal{s}_{t\mathbf{b}}) - \ln(\mathcal{s}_{t0}) = \sum_{j\in\mathbf{b}}(\delta_j + x_{tj}\beta - \alpha p_{tj} + \xi_{tj}) + \ln\left(\sum_{j\in\mathbf{b}} \omega_{\mathbf{b}j}\left(\mathcal{s}_{t(\mathbf{b}|j)}\right)^{1-\lambda_j}\right) + \Gamma_{t\mathbf{b}} \quad (3)$$

with $\omega_{\mathbf{b}j} = \mathbb{1}_{\mathbf{b}\in\mathbf{N}_j} \;/\; \sum_{j'=1}^{J} \mathbb{1}_{\mathbf{b}\in\mathbf{N}_{j'}}$ a weight that proportionally allocates **b** to nests

# PONL Model: Overlapping Nests and Berry (1994)

The PONL inverse demand for **bundle b** in market $t$

$$\ln(s_{t\mathbf{b}}) - \ln(s_{t0}) = \sum_{j\in\mathbf{b}}(\delta_j + x_{tj}\beta - \alpha p_{tj} + \xi_{tj}) + \ln\left(\sum_{j\in\mathbf{b}} \omega_{\mathbf{b}j}\left(s_{t(\mathbf{b}|j)}\right)^{1-\lambda_j}\right) + \Gamma_{t\mathbf{b}} \quad (3)$$

while for **one unit of product $j$** $\implies$ $\mathbf{b} = j$, $\omega_{\mathbf{b}j} = 1$, and $\Gamma_{t\mathbf{b}} = 0$

$$\ln s_{tj} - \ln s_{t0} = \delta_j + x_{tj}\beta - \alpha p_{tj} + (1 - \lambda_j)\ln(s_{t(j|j)}) + \xi_{tj}, \quad (4)$$

# PONL Model: Overlapping Nests and Berry (1994)

The PONL inverse demand for **bundle b** in market $t$

$$\ln(\mathscr{s}_{t\mathbf{b}}) - \ln(\mathscr{s}_{t0}) = \sum_{j\in\mathbf{b}}(\delta_j + x_{tj}\beta - \alpha p_{tj} + \xi_{tj}) + \ln\left(\sum_{j\in\mathbf{b}}\omega_{\mathbf{b}j}\left(\mathscr{s}_{t(\mathbf{b}|j)}\right)^{1-\lambda_j}\right) + \Gamma_{t\mathbf{b}} \quad (3)$$

while for **one unit of product $j$**

$$\ln \mathscr{s}_{tj} - \ln \mathscr{s}_{t0} = \delta_j + x_{tj}\beta - \alpha p_{tj} + (1-\lambda_j)\ln(\mathscr{s}_{t(j|j)}) + \xi_{tj}, \quad (4)$$

- **If $\mathscr{s}_{t(\mathbf{b}|j)}$ observed** $\implies$ **Berry (1994): 2SLS from** (4) **+ plug-in from** (3).

# PONL Model: Overlapping Nests and Berry (1994)

The PONL inverse demand for **bundle b** in market $t$

$$\ln(s_{t\mathbf{b}}) - \ln(s_{t0}) = \sum_{j\in\mathbf{b}}(\delta_j + x_{tj}\beta - \alpha p_{tj} + \xi_{tj}) + \ln\left(\sum_{j\in\mathbf{b}} \omega_{\mathbf{b}j}\left(s_{t(\mathbf{b}|j)}\right)^{1-\lambda_j}\right) + \Gamma_{t\mathbf{b}} \quad (3)$$

while for **one unit of product $j$**

$$\ln s_{tj} - \ln s_{t0} = \delta_j + x_{tj}\beta - \alpha p_{tj} + (1-\lambda_j)\ln(s_{t(j|j)}) + \xi_{tj}, \quad (4)$$

- **If $s_{t(\mathbf{b}|j)}$ observed** $\implies$ **Berry (1994): 2SLS from** (4) **+ plug-in from** (3).
- **Unfortunately, not in PONL**: only observe $s_{t\mathbf{b}} = \sum_{j=1}^{J} s_{t(\mathbf{b}|j)} s_{t\mathbf{N}_j}$. ⏩ Details

# Estimation: Concentrated Two-Stage Least Square (C2SLS)

- Consider each unobserved $\mathcal{s}_{t(j|j)} = \pi_{tj}$ as an additional **auxiliary parameter**.

# Estimation: Concentrated Two-Stage Least Square (C2SLS)

- Consider each unobserved $\mathscr{s}_{t(j|j)} = \pi_{tj}$ as an additional **auxiliary parameter**.

- **Augment 2SLS system with nonliner equations** for $\boldsymbol{\pi}_t = (\pi_{tj})_{j=1}^{J}$.
  - To pin down $\boldsymbol{\pi}_t$, need also to determine demand synergies $\boldsymbol{\Gamma}_t$.
  - Use the remaining equations implied by PONL model.

# Estimation: Concentrated Two-Stage Least Square (C2SLS)

- Consider each unobserved $\mathscr{s}_{t(j|j)} = \pi_{tj}$ as an additional **auxiliary parameter**.

- **Augment 2SLS system with nonliner equations** for $\boldsymbol{\pi}_t = (\pi_{tj})_{j=1}^{J}$.
  - To pin down $\boldsymbol{\pi}_t$, need also to determine demand synergies $\boldsymbol{\Gamma}_t$.
  - Use the remaining equations implied by PONL model.

- **Intuitively**: **Concentrate** out unknown $\boldsymbol{\pi}_t$ from **2SLS** by Berry (1994).

# Implementation: A Convenient Iterative Procedure

- **Properties of Concentrated 2SLS (C2SLS)**:
  - **Good in theory**: PONL identified; C2SLS consistent and asym. normal. Details
  - **Hard in practice**: directly solving nonlinear system complex with large choice sets.

# Implementation: A Convenient Iterative Procedure

- **Properties of Concentrated 2SLS (C2SLS)**:
  - **Good in theory**: PONL identified; C2SLS consistent and asym. normal. Details
  - **Hard in practice**: directly solving nonlinear system complex with large choice sets.

- Use **Gauss-Seidel iterative procedure** to implement C2SLS (Hallett, 1982).

- Algorithm only involves **iterating between linear regressions and plug-ins**.

# Implementation: A Convenient Iterative Procedure

Three **computational advantages**:

(i) Optimization-free.

(ii) Derivative-free.

(iii) Fully parallelizable over $(t, \mathbf{b})$.

# Implementation: A Convenient Iterative Procedure

Three **computational advantages**:

(i) Optimization-free.

(ii) Derivative-free.

(iii) Fully parallelizable over $(t, \mathbf{b})$.

On the **convergence of the algorithm**:

- (Necessary Condition) Convergence of the algorithm $\Longrightarrow$ C2SLS.

# Implementation: A Convenient Iterative Procedure

Three **computational advantages**:

(i) Optimization-free.

(ii) Derivative-free.

(iii) Fully parallelizable over $(t, \mathbf{b})$.

On the **convergence of the algorithm**:

- (Necessary Condition) Convergence of the algorithm $\implies$ C2SLS.
- Numerical convergence (or lack of it) can be easily verified.

# Quantity Discounts in Carbonated Soft Drinks

- **IRI Data**: 6,155 households purchasing 16,873 different bundles of CSDs in Pittsfield and Eau Claire (USA), 2008-2011.
- **Discretize Quantity**: Consider purchases up to 1L as one unit, between 1L and 2L as two units, etc.

# Quantity Discounts in Carbonated Soft Drinks

- **IRI Data**: 6,155 households purchasing 16,873 different bundles of CSDs in Pittsfield and Eau Claire (USA), 2008-2011.
- **Discretize Quantity**: Consider purchases up to 1L as one unit, between 1L and 2L as two units, etc.

**Facts in the data:**

1. Prevalence of **purchases of multiple units** (93.24% shopping trips) of the same and of different products.
2. **Multi-person households purchase larger bundles** than single-person households.
3. **Pervasiveness of quantity discounts**.

# Demand Specification: PONL Model by Household Size

- Average utility of household of size $hs \in \{\text{single, multi}\}$ for **one unit of $j$** and for **bundle b**, respectively:

$$\delta_{tj}^{hs} = \delta_j^{hs} - \alpha^{hs} p_{tj} + \delta_{\text{store}(t)} + \delta_{\text{time}(t)} + \xi_{tj}^{hs}$$

$$\delta_{t\mathbf{b}}^{hs} = \sum_{j \in \mathbf{b}} \delta_{tj}^{hs} + \Gamma_{t\mathbf{b}}^{hs}.$$

- Because of **quantity discounts**, the **≈176,700 demand synergies** are:

$$\Gamma_{t\mathbf{b}}^{hs} = -\alpha^{hs} \left( p_{t\mathbf{b}} - \sum_{j \in \mathbf{b}} p_{tj} \right) + \gamma_{t\mathbf{b}}^{hs}.$$

- Practical implementation of C2SLS:
  - Hausman-type instruments, prices from other cities (Hausman, 1996; Nevo, 2001).
  - **Iterative procedure converges in a couple of minutes (25 iterations)**.

# Demand Estimates

- **Nesting parameter** $\lambda$ similar across household sizes, around 0.88.
- **Price coefficients**: $\alpha^{\text{single}} = 0.75 <^{***} \alpha^{\text{multi}} = 1.03$.

- **Multi-person households more price elastic**.
  - When prices of multiple units increase, multi-person households substitute away from larger bundles more sharply than single-person households.

# Demand Estimates

- **Nesting parameter** $\lambda$ similar across household sizes, around 0.88.
- **Price coefficients**: $\alpha^{\text{single}} = 0.75 <^{***} \alpha^{\text{multi}} = 1.03$.

- **Multi-person households more price elastic**.
  - When prices of multiple units increase, multi-person households substitute away from larger bundles more sharply than single-person households.

- **Incentives for quantity discounts**: demand synergies rather than cost savings.
  - Estimated **marginal costs non-decreasing** in quantities.

# Counterfactual Linear Pricing

1. **Profitable for producers of CSDs** in line with Varian (1992).
   - From quantity discounts to linear pricing, industry profit down by 19.7%.
2. **Large reduction in purchased quantities** ($-20.7\%$).
   - Price increase for larger quantities ($+14.9\%$).
   - Price decrease for smaller quantities ($-31.6\%$).
3. **Consumer surplus** remains small.
   - CV of $+3.7\$$ per household-year (2.8% of total expenditure of CSDs).
   - Contraction in purchased quantities, but also lower prices for single units.

# Counterfactual Linear Pricing

1. **Profitable for producers of CSDs** in line with Varian (1992).
   - From quantity discounts to linear pricing, industry profit down by 19.7%.
2. **Large reduction in purchased quantities** ($-20.7\%$).
   - Price increase for larger quantities ($+14.9\%$).
   - Price decrease for smaller quantities ($-31.6\%$).
3. **Consumer surplus** remains small.
   - CV of $+3.7\$$ per household-year (2.8% of total expenditure of CSDs).
   - Contraction in purchased quantities, but also lower prices for single units.

**Policy question**: Could a ban on quantity discounts serve as a policy to limiting added sugar intake from CSDs?

# Counterfactual Linear Pricing: Changes in Added Sugar Intake

| | Ban on quantity discounts on all CSDs | Ban on quantity discounts only on sugary CSDs | Sugar tax 1¢/oz of added sugar |
|---|---|---|---|
| **Predicted added sugar change** | −22.93% | −22.08% | −22.90% |
| **Quantity change** | −20.66% | −8.93% | −10.35% |
| Sugary CSDs | −23.95% | −21.89% | −21.98% |
| Non-Sugary CSDs | −17.83% | +1.71% | +2.14% |
| **Profit change** | −19.74% | −9.46% | −7.01% |
| **CV** ($ per household-year) | +3.70$ | +1.77$ | +2.35$ |
| **CV/Expenditure** | +2.82% | +1.29% | +1.61% |

# Conclusions

- Propose **empirical model of demand for bundles**:
  1. Accommodates intuitive form of correlation in the preferences of bundles.
  2. **Convenient** in applications **with large choice sets**.

- Inform policy debate on a **ban on quantity discounts in CSDs**.

# Conclusions

- Propose **empirical model of demand for bundles**:
  1. Accommodates intuitive form of correlation in the preferences of bundles.
  2. **Convenient** in applications **with large choice sets**.

- Inform policy debate on a **ban on quantity discounts in CSDs**.

- **Proposed model can also facilitate the study of**:
  - Demand across multiple product categories (grocery, online shopping, etc.).
  - Mergers in markets with both substitutes and complements.
  - Spillovers of taxes from a product category to others.
  - Portfolio choice models of asset pricing.
  - ...

# Backup Slides

# Lack of Observability of Within-Nest Purchase Probabilities ⏩ Back

- Because some **b** $\in$ **multiple nests**, cannot determine $\mathcal{s}_{t(\mathbf{b}|j)}$'s from $\mathcal{s}_{t\mathbf{b}}$'s.

- $\mathbf{N}_1 = \{1, (1,1), (1,2), (1,3)\}$, $\mathbf{N}_2 = \{2, (2,2), (1,2)\}$, and $\mathbf{N}_3 = \{3, (1,3)\}$.

- Then **8 observed purchase probabilities** and **9 unknowns**:

$$
\begin{aligned}
\mathcal{s}_{tk} &= \mathcal{s}_{t(k|k)} \mathcal{s}_{t\mathbf{N}_k} \quad k = 1,2,3 \\
\mathcal{s}_{t(j,j)} &= \mathcal{s}_{t(j,j|j)} \mathcal{s}_{t\mathbf{N}_j} \quad j = 1,2 \\
\mathcal{s}_{t(1,2)} &= \mathcal{s}_{t(1,2|1)} \mathcal{s}_{t\mathbf{N}_1} + \left(1 - \mathcal{s}_{t(2|2)} - \mathcal{s}_{t(2,2|2)}\right) \mathcal{s}_{t\mathbf{N}_2} \\
\mathcal{s}_{t(1,3)} &= \left(1 - \mathcal{s}_{t(1|1)} - \mathcal{s}_{t(1,1|1)} - \mathcal{s}_{t(1,2|1)}\right) \mathcal{s}_{t\mathbf{N}_1} + \left(1 - \mathcal{s}_{t(3|3)}\right) \mathcal{s}_{t\mathbf{N}_3} \\
\mathcal{s}_{t0} &= 1 - \sum_{j=1}^{3} \mathcal{s}_{t\mathbf{N}_j}.
\end{aligned}
\tag{5}
$$

# Details: Estimation ⏩ Back

(i) **Uniqueness** of $\Gamma_t$ and $\pi_t$ given $\lambda$ for each market $t$.

- $\pi_t$ uniquely determined as $\pi_t = \pi(\lambda; \mathcal{s}_t)$.
- $\Gamma_{t\mathbf{b}}$ uniquely determined as $\Gamma_{t\mathbf{b}} = \Gamma_{\mathbf{b}}(\lambda; \pi_t, \mathcal{s}_t) = \Gamma_{\mathbf{b}}(\lambda; \pi(\lambda; \mathcal{s}_t), \mathcal{s}_t)$.
- This means that we can **concentrate out** $\Gamma_t$ and $\pi_t$ in each $t$.

(ii) Given (i), C2SLS reduces to a nonlinear system in $(\delta, \beta, \alpha, \lambda)$.

- Assume **rank condition** at the true parameters of the nonlinear system. Two roles:
- **First**, *in finite samples* existence of $(\hat{\delta}, \hat{\beta}, \hat{\alpha}, \hat{\lambda})$ in neighbourhood of true values with probability one as $T \to \infty$.
- **Second**, *asymptotically* $(\hat{\delta}, \hat{\beta}, \hat{\alpha}, \hat{\lambda})$ consistent and normal.

(iii) Finally, $\hat{\Gamma}_t$ and $\hat{\pi}_t$ also consistent and asymptotically normal as functions of $(\hat{\delta}, \hat{\beta}, \hat{\alpha}, \hat{\lambda})$.

# Choice of Instruments ⏩ Back

- Despite the **lack of observability of $\pi_{tj}$**, instruments for $\frac{\mathcal{s}_{tj}/\mathcal{s}_{t0}}{\pi_{tj}}$ can be chosen on the basis of their correlation with $\mathcal{s}_{t(j|j)}$.

- Denote by $\pi_j(\lambda; \mathcal{s}_t)$ the unique $\pi_{tj}$ that rationalizes the model for given $(\lambda; \mathcal{s}_t)$.

- 1st-order Taylor approx. of $\ln(\pi_j(\lambda; \mathcal{s}_t))$ around true value $\ln(\pi_j(\lambda^0; \mathcal{s}_t))$:

$$
\begin{aligned}
\ln \mathcal{s}_{tj} - \ln \mathcal{s}_{t0} &= \delta_j + x_{tj}\beta - \alpha p_{tj} + (1-\lambda_j)\left[\ln\left(\frac{\mathcal{s}_{tj}/\mathcal{s}_{t0}}{\pi^0_{tj}}\right) - \frac{1}{\mathcal{s}^j_t}\frac{\partial \pi_j(\lambda^0; \mathcal{s}_t)}{\partial \lambda}(\lambda - \lambda^0)\right] + \xi_{tj} \\
&= \delta_j + x_{tj}\beta - \alpha p_{tj} + (1-\lambda_j)\left[\ln \mathcal{s}_{t(j|j)} - \frac{1}{\mathcal{s}^j_t}\frac{\partial \pi_j(\lambda^0; \mathcal{s}_t)}{\partial \lambda}(\lambda - \lambda^0)\right] + \xi_{tj}
\end{aligned}
$$

- $\ln \mathcal{s}_{t(j|j)}$ **leading term**: **valid IVs "shift"** $\ln \mathcal{s}_{t(j|j)}$ **independently of** $\xi_{tj}$.

# Choice of Instruments ⏩ Back

Re-express $\mathcal{s}_{t(j|j)}$ as:

$$\mathcal{s}_{t(j|j)} = \frac{\exp(\delta_{tj})^{1/\lambda_j}}{\sum_{\mathbf{b}' \in \mathbf{N}_j} (\omega_{\mathbf{b}'j} \exp(\delta_{t\mathbf{b}'}))^{1/\lambda_j}} = \frac{1}{1 + \sum_{\mathbf{b}' \in \mathbf{N}_j, \mathbf{b}' \neq j} (\omega_{\mathbf{b}'j} \exp(\delta_{t\mathbf{b}'} - \delta_{tj}))^{1/\lambda_j}}.$$

- **Valid IVs shift $\delta_{t\mathbf{b}'} - \delta_{tj}$ independently of $\xi_{tj}$ for $\mathbf{b}' \in \mathbf{N}_j$.**
- **Differentiation IVs** (Gandhi and Houde, 2019)
    - $x_{t\mathbf{b}'} - x_{tj}$.
    - $x_{tk} - x_{tj}$ for $k \neq j$ as long as nests $j$ and $k$ are overlapping, $\mathbf{N}_k \cap \mathbf{N}_j \neq \emptyset$.
- **Excluded prices** as $p_k$ for $k \neq j$ and $\mathbf{N}_k \cap \mathbf{N}_j \neq \emptyset$
    - Cost shifters (or their proxies) for $p_k$.

# References


Anderson, S. P. and Renault, R. (2011). *Price discrimination*. Edward Elgar Publishing.

Armstrong, M. (2016). Nonlinear pricing. *Annual Review of Economics*, 8:583–614.

Berry, S., Khwaja, A., Kumar, V., Musalem, A., Wilbur, K. C., Allenby, G., Anand, B., Chintagunta, P., Hanemann, W. M., Jeziorski, P., et al. (2014). Structural models of complementary choices. *Marketing Letters*, 25(3):245–256.

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, pages 242–262.

Compiani, G. (2019). Market counterfactuals and the specification of multi-product demand: A nonparametric approach. *Working Paper*.

Dubé, J.-P. (2004). Multiple discreteness and product differentiation: Demand for carbonated soft drinks. *Marketing Science*, 23(1):66–81.

Gandhi, A. and Houde, J.-F. (2019). Measuring substitution patterns in differentiated products industries. Technical report, National Bureau of Economic Research.

Gentzkow, M. (2007). Valuing new goods in a model with complementarity: Online newspapers. *The American Economic Review*, 97(3):713–744.

Hallett, A. H. (1982). Alternative techniques for solving systems of nonlinear