Estimating Discrete-Choice Models of Product Differentiation

Steven Berry (1994), *RAND Journal of Economics*

Chenhui Lu

University of Florida

Sep. 24 2025

Background: General Problem

- In differentiated product markets, reliable estimation of **demand and costs** is challenging.
- Prices are typically **correlated with unobserved product characteristics** (quality, brand perception, style) $\Rightarrow$ **endogeneity**.
- Ignoring endogeneity yields biased estimates (e.g., spurious findings that higher prices increase demand).
- **Societal relevance**: accurate demand is essential for antitrust, merger policy, and welfare analysis.

Specific Research Problem

- How to model consumer choice with **product differentiation** in a discrete-choice framework.
- How to address **price endogeneity** in demand estimation.

Why Is It Important?

- In IO, Public, **demand elasticities** are central for policy and welfare analysis.
- Traditional methods for homogeneous goods fail in differentiated settings.
- This paper introduces a method that became the foundation for the **BLP framework** (Berry, Levinsohn, Pakes 1995).

Consumer Demand Model

- Consumers choose among N products and an outside good.
- Indirect utility:

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

- Components: x_j observed characteristics; p_j price ($\alpha > 0$); ξ_j unobserved quality; ϵ_{ij} error term.
- Each consumer chooses the product with the highest utility.

The Endogeneity Problem

- Unobserved quality ξ_j affects both **demand** and **price**.
- Better products tend to be priced higher $\Rightarrow$ p_j and ξ_j are correlated.
- Direct regressions of market share on price are biased (sometimes suggesting consumers prefer higher prices).
- **Key challenge**: separate the true effect of price from unobserved quality.

1 Introduction

2 Methodology

Consumer Demand and Endogeneity

Inversion and IV Estimation

Special Cases of Inversion

Supply and Joint Estimation

3 Results

4 Contributions

5 Discussion

Innovation 1: Mean Utility Inversion

- Market shares s_j map **one-to-one** to mean utilities δ_j.
- Define mean utility:

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

- **Inversion** recovers δ_j from observed shares.
- Special cases:
  - **Logit**: $\delta_j = \ln(s_j) - \ln(s_0)$
  - **Nested logit**: includes a within-group term
- Interpretation: shares contain information on product attractiveness; inversion recovers latent utility.

Innovation 2: IV Estimation

- With δ_j in hand, estimate

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

- Endogeneity remains: p_j correlated with ξ_j.
- **Instrumental variables (IV)**:
  - Cost shifters (input prices, wages, transportation costs)
  - Rival product characteristics (affect equilibrium pricing but not utility directly)

Special Cases of Inversion

Logit model

$$\ln(s_j) - \ln(s_0) = x_j\beta - \alpha p_j + \xi_j$$

Simplest closed-form inversion.

Vertical differentiation

- Products are ordered by quality; market shares defined by cutoff rules.

Nested logit

- Allows correlation within groups; adds $\sigma \ln(s_{j|g})$.

Random coefficients logit

- Most general: consumer heterogeneity interacts with characteristics.
- Market shares approximated via simulation.

1. Introduction
2. Methodology
   - Consumer Demand and Endogeneity
   - Inversion and IV Estimation
   - Special Cases of Inversion
   - Supply and Joint Estimation
3. Results
4. Contributions
5. Discussion

Supply Side (Firm Pricing)

- Assume **Bertrand Nash** competition.
- First-order condition (FOC):

$$p_j = c_j + \text{markup}_j$$

- Markups depend on demand elasticities:
  - Few substitutes $\Rightarrow$ higher markups
  - Intense competition $\Rightarrow$ lower markups

Joint Estimation Framework

Steps

1. Invert observed shares to recover δ_j.
2. Estimate demand parameters (β, α) via IV.
3. Use supply-side pricing FOC to identify cost parameters.
4. **Joint GMM**: combine demand and supply moments to estimate the full parameter vector.
5. Outcome: preference and cost estimates that enable **counterfactual policy analysis**.

Advantages of the Approach

Flow

(Observed shares s_j) $\rightarrow$ [Inversion: $\delta_j(s)$] $\rightarrow$ [IV Regression $\Rightarrow$ β, α]
Demand moments: $\delta_j = x_j\beta - \alpha p_j + \xi_j$ Supply moments: $p_j = c_j(\gamma) + \text{markup}$
$\Downarrow$ Joint GMM Estimation $\Rightarrow$ Demand & Cost Parameters $\Rightarrow$ Counterfactuals

Monte Carlo Experiments

- Simulation of 500 duopoly markets.

TABLE 1 Monte Carlo Parameter Estimates 100 Random Samples of 500 Duopoly Markets Logit Utility

		$(\sigma_{\xi d} = 1)$		$(\sigma_{\xi d} = 3)$	
Parameter	True Value	(1) OLS	(2) IV	(3) OLS	(4) IV
β_o	5	3.46	4.98	0.378	4.89
		(.158)	(.226)	(.415)	(.738)
β_x	2	1.41	1.99	.325	1.95
		(.058)	(.091)	(.127)	(.272)
α	1	.726	.995	.181	.979
		(.029)	(.039)	(.076)	(.128)

Notes: The values given in the table are empirical means and (standard errors).
The utility function is $u_{ij} = \beta_o + \beta_x x_j + \sigma_{\xi d}\xi_j - \alpha p_j + \epsilon_{ij}$.
Marginal cost is $c_j = e^{\gamma_o + \gamma_x x_j + \sigma_{\xi c}\xi_j + \gamma_w w_j + \sigma_\omega \omega_j}$.

- **OLS**: severely underestimates the price coefficient; may imply consumers prefer higher prices.
- **IV**: recovers true parameters consistently.
- Demonstrates that ignoring unobserved characteristics yields **systematic bias**.

Empirical Applications

- **Automobile industry** (Berry, Levinsohn, Pakes 1993, 1995): plausible demand elasticities and substitution patterns.
- **Computer industry** (Greenstein 1992): vertical differentiation with sensible estimates.
- The framework has become a **standard tool** in empirical IO.

Contributions

1. Introduces **mean-utility inversion** with **IV** estimation.
2. Provides a solution to **price endogeneity** in differentiated demand.
3. Integrates demand with **supply-side FOC**.

Discussion Questions

- Why can market shares be inverted to recover mean utilities?
- Why are rival product characteristics valid instruments?
- Why does the random coefficients logit require simulation?

Thank You