Demand Estimation 5
PhD Industrial Organization

Nicholas Vreugdenhil

Plan

1. Distinguishing models of competition
2. Supply-side moments
3. Welfare from new products: theory
4. Welfare from new products: application
5. Apple-cinnamon cheerios war
6. Main takeaways of demand estimation

Plan

1. **Distinguishing models of competition**
2. Supply-side moments
3. Welfare from new products: theory
4. Welfare from new products: application
5. Apple-cinnamon cheerios war
6. Main takeaways of demand estimation

Application: distinguishing between models of competition

- **Overall strategy:**
- 1. Estimate demand
- 2. Use estimates + pricing rules implied by different models of firm conduct to get price-cost margins (PCM)
  - Challenge: costs not observed
- 3. Compare predicted PCM from different models of conduct to true PCM
  - See which model of firm conduct best matches the data.
- **Main finding:**
- Nash-Bertrand pricing best matches observed PCM

Application: distinguishing between models of competition

- **Demand side:**

- Utility model is exactly what we have seen:

$$u_{ijt} = x_{jt}\beta_i + \alpha_i p_{jt} + \xi_{jt} + \epsilon_{ijt}$$

- **Data:**

- Use top 25 cereal brands

- Scanner data $\rightarrow$ get market shares, prices, etc.
  - Aggregate to MSA-quarter level = 1124 markets.

- Advertising data

- Cereal box characteristics (nutritional information); subjective characteristics ('mushy')

- Demographics from CPS

Application: distinguishing between models of competition

- **Supply side:**

- Profits of firm f:

$$\pi_f = \sum_{j \in \mathcal{J}_f} \left[(p_j - mc_j) q_j(\mathbf{p}) - FC_j \right]$$

- Where:
  - π_f: profits of firm f
  - p_j: price of product j; $\mathbf{p}$: vector of prices
  - mc_j: marginal cost
  - FC_j: fixed cost
  - q_j: quantity of product j (depends on <u>all</u> prices)
  - $\mathcal{J}_f$: set of products that firm f maximizes profit over

Application: distinguishing between models of competition

- **Supply side:**

- Define **conduct structure** as JxJ matrix:

$$H_{jk} = \begin{cases} 1 & \text{if } \exists \text{f where } \{j, k\} \subset \mathcal{J}_f \\ 0 & \textit{otherwise} \end{cases}$$

- Elements of H either 0 or 1

- If value of element $=$ 1: then product j and k are priced *as if jointly owned*

- Examples:
  - Single product firm pricing: identity matrix

  - Joint profit maximization: matrix of 1's

Application: distinguishing between models of competition

- **Supply side:**

- Define $\Omega_{jk} = -\partial q_k / \partial p_j \cdot H_{jk}$
    - Recall: j is index, k is column

- First order condition of firms' profit maximization problem (bold denotes vectors):

$$\mathbf{q}(\mathbf{p}) - \Omega(\mathbf{p} - \mathbf{mc}) = 0$$

- Implies pricing equation:

$$\mathbf{p} - \mathbf{mc} = \Omega^{-1}\mathbf{q}(\mathbf{p})$$

- Important: above equation implies that given conduct structure + estimates of demand substitution $\Omega \rightarrow$ can measure price-cost margins *without observing cost data*.

Application: distinguishing between models of competition: estimation

- Only estimate demand
- Use BLP algorithm
- Instruments: Hausman instruments (prices in other markets)

TABLE VI

RESULTS FROM THE FULL MODEL[a]

Variable	Means (β's)	Standard Deviations (σ's)	Interactions with Demographic Variables: Income	Income Sq	Age	Child
Price	−27.198	2.453	315.894	−18.200	—	7.634
	(5.248)	(2.978)	(110.385)	(5.914)		(2.238)
Advertising	0.020	—	—	—	—	—
	(0.005)					
Constant	−3.592[b]	0.330	5.482	—	0.204	—
	(0.138)	(0.609)	(1.504)		(0.341)	
Cal from Fat	1.146[b]	1.624	—	—	—	—
	(0.128)	(2.809)				
Sugar	5.742[b]	1.661	−24.931	—	5.105	—
	(0.581)	(5.866)	(9.167)		(3.418)	
Mushy	−0.565[b]	0.244	1.265	—	0.809	—
	(0.052)	(0.623)	(0.737)		(0.385)	
Fiber	1.627[b]	0.195	—	—	—	−0.110
	(0.263)	(3.541)				(0.0513)
All-family	0.781[b]	0.1330	—	—	—	
	(0.075)	(1.365)				
Kids	1.021[b]	2.031	—	—	—	
	(0.168)	(0.448)				
Adults	1.972[b]	0.247	—	—	—	
	(0.186)	(1.636)				
GMM Objective (degrees of freedom)			5.05 (8)			
MD χ^2			3472.3			
% of Price Coefficients > 0			0.7			

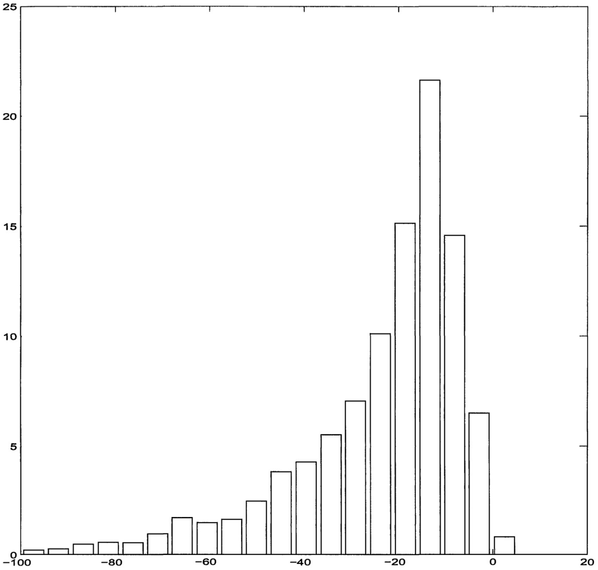

FIGURE 2.—Frequency distribution of price coefficient (based on Table VI).

TABLE VIII

MEDIAN MARGINS[a]

	Logit (Table V column ix)	Full Model (Table VI)
Single Product Firms	33.6%	35.8%
	(31.8%–35.6%)	(24.4%–46.4%)
Current Ownership of 25 Brands	35.8%	42.2%
	(33.9%–38.0%)	(29.1%–55.8%)
Joint Ownership of 25 Brands	41.9%	72.6%
	(39.7%–44.4%)	(62.2%–97.2%)

Application: distinguishing between models of competition: mergers

- Common use of the framework in this paper can also be used to determine the effects of a merger. How would you do this?

Application: distinguishing between models of competition: mergers

- Common use of the framework in this paper can also be used to determine the effects of a merger. How would you do this?
- <u>Answer:</u>
- 1. Using pre-merger data estimate demand and recover marginal costs by inverting the pricing equation:

$$\mathbf{mc} = \mathbf{p} - \Omega^{-1}\mathbf{q}(\mathbf{p})$$

- 2. Change the conduct structure H so that the merging firms jointly maximize the profits of their products

Plan

1. **Supply-side moments**
2. Welfare from new products: theory
3. Welfare from new products: application
4. Apple-cinnamon cheerios war
5. Main takeaways of demand estimation

Supply-side moments

- Previously, we used our demand model + different assumptions about how firms set prices to 'test' models of firm conduct.
- In a related idea, we could alternatively:
  - make assumptions about how firms choose prices
  - utilize data on the supply side to help identify the demand model

Supply-side moments

- Assume marginal cost is given by:

$$mc_{jt} = w_{jt}\gamma + \omega_{jt}$$

- Where:
  - w_{jt}: vector of observed characteristics of product j
  - ω_{jt}: unobserved component
  - γ: parameters to be estimated

Supply-side moments

- Also assume Nash-Bertrand pricing model and combine with cost parametrization on previous slide (using the notation from the previous lecture)

$$\mathbf{p}_t = w_t \gamma + \Omega^{-1} \mathbf{q}(\mathbf{p}_t) + \boldsymbol{\omega}_t$$

- Can form supply-side moments: $E(\omega_{jt}|\mathbf{Z}_t) = 0$.
- Here: $\mathbf{Z}_t$: vector of IVs that include product characteristics, and cost shifters
- Note that above equation is informative about both supply parameters γ and the demand parameters (which affect Ω)
- Can include these additional moments in the GMM step

Plan

1. Supply-side moments
2. **Welfare from new products: theory**
3. Welfare from new products: application
4. Apple-cinnamon cheerios war
5. Main takeaways of demand estimation

Consumer welfare

- Define the **consumer surplus from a logit model**
- Expected utility prior to observing the i.i.d. logit draws from $\{1, 2, \dots, J\}$ choice alternatives: (the **'inclusive value'**)

$$\omega_{it} = E_{\{\epsilon_{i0t},\dots,\epsilon_{iJt}\}} \max_j \{\delta_{jt} + \mu_{ijt} + \epsilon_{ijt}\} = \ln\left(\sum_j \exp\{\delta_{jt} + \mu_{ijt}\}\right)$$

- This formula is also known as the 'log-sum' formula.
- If utility is linear in price, inclusive value can be converted into dollars by dividing by the price coefficient.
- You have (probably) seen this value before, since it comes up when computing a nested logit

Consumer welfare

- Typically, two cases where we compute welfare.
- Case 1: observe quantities and prices and want to summarize them into a welfare measure.
  - Key issue: before we normalized the utility of the outside option to zero.
  - This is fine for estimating choice probabilities (why?)
  - But, issues occur if want to compute inclusive value over time (or across markets)
  - This is because we would be *implicitly assuming that utility from the outside good is constant over time*.

Consumer welfare

- Typically, two cases where we compute welfare.
- Case 1: observe quantities and prices and want to summarize them into a welfare measure.
  - Example: What if we see the share of the 'inside' products increasing over time.
  - Question: What could cause this?

Consumer welfare

- Typically, two cases where we compute welfare.
- Case 1: observe quantities and prices and want to summarize them into a welfare measure.
    - Example: What if we see the share of the 'inside' products increasing over time.
    - Question: What could cause this?
    - Answer: price of inside goods decreased (or quality ξ_{jt} increased) OR the outside option got worse. Different welfare implications, but assuming outside good is 0 rules out latter.
    - Partial solution: Nevo (2003) compute welfare over time for when ξ_{jt} changes / outside option constant vs ξ_{jt} fixed /outside option flexible. Report both extreme cases.

Consumer welfare

- Case 2: Use the model to compute a welfare gain from a counterfactual outcome.

- Assume we observe one market over time (denoted by t).

- Can show that the change in welfare from introducing a product in period t (which comes from the change in the inclusive value) is:

$$\textbf{Logit:}\quad \ln\left(\frac{1}{s_{0t}}\right) - \ln\left(\frac{1}{s_{0t-1}}\right)$$

$$\textbf{Mixed logit:}\quad \int \ln\left(\frac{1}{s_{i0t}}\right) dF(D_{it}, \nu_{it}) - \int \ln\left(\frac{1}{s_{0t-1}}\right) dF(D_{it-1}, \nu_{it-1})$$

- So, logit model: welfare directly related to share of the outside good.

- Mixed logit: same idea, but difference depends on heterogeneity of choosing the outside option.

Consumer welfare

- Case 2: Use the model to compute a welfare gain from a counterfactual outcome.
- Problem with computing welfare from new goods: **red-bus blue-bus problem**.
- Thought experiment:
- Market where consumers choose how to commute.
  - Choices: Car, Red Bus
  - Assume half consumers choose Car, half choose Red Bus
- Assume we artificially introduce a new option: **the Blue Bus**
  - Artificial because we also **assume consumers are color-blind**
  - (Also assume price, frequency of service etc are not impacted)
  - Now, half consumers choose a car, and the rest are split between the two buses.
  - Clearly, **consumer welfare has not changed**.

Consumer welfare

- Case 2: Use the model to compute a welfare gain from a counterfactual outcome.
- **What if we now use our logit model to estimate the welfare effects of introducing a Blue Bus?**
  - (Suppose we only observe data pre-introduction of the Blue Bus.)

Consumer welfare

- Case 2: Use the model to compute a welfare gain from a counterfactual outcome.
- **What if we now use our logit model to estimate the welfare effects of introducing a Blue Bus?**
    - (Suppose we only observe data pre-introduction of the Blue Bus.)
    - Pre-introduction:
    - Assume car is outside good, normalize to zero: $\delta_{car} = 0$. Then, also, $\delta_{\text{red-bus}} = 0$ since $s_{car} = s_{\text{red-bus}} = 0$.
    - Inclusive value is: $\ln(e^0 + e^0) = \ln(2)$.
    - Post-introduction:
    - $\delta_{\text{blue-bus}} = \delta_{\text{red-bus}} = 0$ (since same bus)
    - So, $s_{car} = s_{\text{blue-bus}} = s_{\text{red-bus}} = 1/3$. Inclusive value is: $\ln(3)$, a **welfare increase**!

Consumer welfare

- Case 2: Use the model to compute a welfare gain from a counterfactual outcome.

- Where did we go wrong with the previous analysis?

Consumer welfare

- Case 2: Use the model to compute a welfare gain from a counterfactual outcome.
- Where did we go wrong with the previous analysis?
- Main issue: we are getting an extra logit draw when we introduce the Blue Bus.
- Possible solution: if we observe the market post-introduction of the new product (i.e. the Blue Bus).
- Post-introduction:
- $s_{car} = 0.5$, $s_{\text{blue-bus}} = s_{\text{red-bus}} = 0.25$
- Implies: $\delta_{\text{blue-bus}} = \delta_{\text{red-bus}} = \ln(0.5)$
- Inclusive value: $\ln(e^0 + 2 * e^{ln(0.5)}) = \ln(2)$. Note that this is the correct answer.

Consumer welfare

- Main takeaways from the above ‘red-bus blue-bus’ exercise:
- 1. Introducing new products also introduces with extra logit draws, which can bias welfare computations
- 2. Observing post-introduction market shares can ‘correct’ for this bias.
- Note that the above is true as well in Mixed Logit and other models:
  - Berry and Pakes (2007): “the fact that the contraction fits the shares exactly means that the extra gain from the logit errors is offset by lower δ’s, and this roughly counteracts the problems generated for welfare measurement by the model with tastes for products.”

Plan

1. Supply-side moments
2. Welfare from new products: theory
3. **Welfare from new products: application**
4. Apple-cinnamon cheerios war
5. Main takeaways of demand estimation

Application of welfare from new products: the minivan (Petrin, 2002)

- **Question:** what are the welfare benefits of the minivan?
- Minivan:
- Box-like vehicle first introduced in 1984
- Increased height, decreased floor, making entry easier and allowing more movement in the vehicle
- Extremely popular with families in the 80s, 90s, etc

Application of welfare from new products: the minivan (Petrin, 2002)

- Motivation: new products, and product improvements, are a key source of economic growth. But how should we actually value these innovations?
- Main takeaways (beyond valuing the minivan):
- 1. Logit model performs extremely poorly compared to mixed logit
- 2. Use of micro-moments
- 3. Dealing with the 'red-bus blue-bus' problem

Application of welfare from new products: the minivan (Petrin, 2002)

- (Conditional, indirect) utility:

$$u_{ijt} = x_{jt}\beta_{it} + \alpha_{it}\ln(y_i - p_{jt}) + \xi_{jt} + \epsilon_{ijt}$$

- Price coefficient allowed to vary by income
- Other notes: ‘minivan’ enters *as a characteristic*, also includes a supply side
- Data
- Observes US market 1981-1993 (for 916 vehicles)
- Prices, quantities (from Automotive Data Book)
- Product characteristics: fuel efficiency, vehicle dimensions, etc
- Important: also has CEX auto supplement, links demographics $\leftrightarrow$ new vehicle purchases

Application of welfare from new products: the minivan (Petrin, 2002)

- Main difference to BLP: includes micro-moments using his data from CEX
  - E[i buys new vehicle | low income]
  - E[i buys new vehicle | mid income]
  - E[i buys new vehicle | high income]
  - E[family size of i | i purchase minivan]
  - E[family size of i | i purchase station wagon]
  - etc...
- Include these as additional moments in the GMM part of the BLP estimator
- Benefit: more info on heterogeneity $\rightarrow$ can better capture substitution patterns
- Advice: add micro-moments where possible
- (Note: also uses product characteristics as instruments)

Application of welfare from new products: the minivan (Petrin, 2002)

TABLE 4
PARAMETER ESTIMATES FOR THE DEMAND-SIDE EQUATION

Variable	OLS Logit (1)	Instrumental Variable Logit (2)	Random Coefficients (3)	Random Coefficients and Microdata (4)
		A. Price Coefficients (α's)		
α_1	.07	.13	4.92	7.52
	(.01)**	(.01)**	(9.78)	(1.24)**
α_2			11.89	31.13
			(21.41)	(4.07)**
α_3			37.92	34.49
			(18.64)**	(2.56)**

Application of welfare from new products: the minivan (Petrin, 2002)

	B. Base Coefficients (β's)			
Constant	−10.03	−10.04	−12.74	−15.67
	(.32)**	(.34)**	(5.65)**	(4.39)**
Horsepower/weight	1.48	3.78	3.40	−2.83
	(.34)**	(.44)**	(39.79)	(8.16)
Size	3.17	3.25	4.60	4.80
	(.26)**	(.27)**	(24.64)	(3.57)*
Air conditioning standard	−.20	.21	−1.97	3.88
	(.06)**	(.08)**	(2.23)	(2.21)*
Miles/dollar	.18	.05	−.54	−15.79
	(.06)**	(.07)	(3.40)	(.87)**
Front wheel drive	.32	.15	−5.24	−12.32
	(.05)**	(.06)**	(3.09)	(2.36)**
Minivan	.09	−.10	−4.34	−5.65
	(.14)	(.15)	(13.16)	(.68)**
Station wagon	−1.12	−1.12	−20.52	−1.31
	(.06)**	(.07)**	(36.17)	(.36)**
Sport-utility	−.41	−.61	−3.10	−4.38
	(.09)**	(.10)**	(10.76)	(.41)**
Full-size van	−1.73	−1.89	−28.54	−5.26
	(.16)**	(.17)**	(235.51)	(1.30)**
% change GNP	.03	.03	.08	.24
	(.01)**	(.01)**	(.02)**	(.02)**

Application of welfare from new products: the minivan (Petrin, 2002)

- Counterfactual exercise: eliminate minivan
- Compute change in consumer welfare ('compensating variation')
- (Note: prices allowed to readjust too using the supply-side)

Application of welfare from new products: the minivan (Petrin, 2002)

TABLE 8
AVERAGE COMPENSATING VARIATION CONDITIONAL ON MINIVAN PURCHASE, 1984: 1982–84 CPI-ADJUSTED DOLLARS

	OLS Logit	Instrumental Variable Logit	Random Coefficients	Random Coefficients and Microdata
Compensating variation:				
Median	9,573	5,130	1,217	783
Mean	13,652	7,414	3,171	1,247

Application of welfare from new products: the minivan (Petrin, 2002)

- CV in logit model is biased upwards (due to large logit draws - i.e. consumers in the model with extreme tastes for minivans)
- Petrin claims: adding the micro-data reduces reliance of the model on the logit draws, and so reduces the overall welfare gain.
- Additionally: directly observes the counterfactual new product choice (so market shares with the new product are correct - we mentioned earlier this was useful in reducing the red-bus blue-bus problem).

Plan

1. Supply-side moments
2. Welfare from new products: theory
3. Welfare from new products: application
4. **Apple-cinnamon cheerios war**
5. Main takeaways of demand estimation

Apple-cinnamon cheerios war

- Hausman paper in "The Economics of New Goods" (1997)
- Values welfare contribution of Apple Cinnamon Cheerios at 60 million dollars per year (in mid 1990s)
- Bresnahan (1997) disagrees with this computation. Criticises identifying assumptions, assumptions about competition, etc...
- Not time to go into it in detail, but on reading list and highly recommended
- Quote from Bresnahan (1997): *'I have never met an economist who strayed that far from reality by ideology – only by arrogance.'*

Plan

1. Supply-side moments
2. Welfare from new products: theory
3. Welfare from new products: application
4. Apple-cinnamon cheerios war
5. **Main takeaways of demand estimation**

Main takeaways of demand estimation part of this course

- We learned the Mixed Logit (‘BLP’) model, a model of **demand for differentiated products**.
- **Main components**: product characteristics, observed demographics, unobserved tastes for characteristics, unobserved demand shocks
- Captures **more consumer heterogeneity** than standard logit. An example where this really matters is getting substitution patterns right.
- We saw how to estimate it using the BLP method, common data that are useful or required, and common computational problems
- We studied common instruments (and discussed identification - e.g. recall **dual role** of instruments)
- We saw some applications: measuring conduct, valuing new goods (there are *many* more - often used whenever you need demand in a model)

Main takeaways of demand estimation part of this course

- Computational note: I will get you to code up a simple version of BLP in the homework.
- However, if you actually use it, I highly recommend the PyBLP package (Conlon and Gortmaker). This will save potentially years of computation/coding time!