Lecture 2: Demand Estimation (Part one)

Chris Conlon

September 3, 2019

Grad IO

Intro

Differentiated Products Bertrand

Consider the multi-product Bertrand FOCs:

$$\begin{aligned}\arg\max_{p \in \mathcal{J}_f} \pi_f(\mathbf{p}) &= \sum_{j \in \mathcal{J}_f} (p_j - c_j) \cdot q_j(\mathbf{p}) \\ \rightarrow 0 &= q_j(\mathbf{p}) + \sum_{k \in \mathcal{J}_f} (p_k - c_k) \frac{\partial q_k}{\partial p_j}(\mathbf{p})\end{aligned}$$

It is helpful to define the matrix Ω with entries:

$$\Omega_{(j,k)}(\mathbf{p}) = \left\{ \begin{array}{cc} -\frac{\partial q_j}{\partial p_k}(\mathbf{p}) & \text{for } (j,k) \in \mathcal{J}_f \\ 0 & \text{for } (j,k) \notin \mathcal{J}_f \end{array} \right\}$$

We can re-write the FOC in matrix form:

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

What do we want to learn from demand systems?

We can recover markups and marginal costs:

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

- We need two objects from the demand system:
  - The (vector of) predicted sales at the (vector of) prices $\mathbf{q}(\mathbf{p})$.
  - The derivatives (or elasticities) of the demand function $\frac{\partial q_k}{\partial p_j}(\mathbf{p})$.
- Ideally, these objects depend on the full vector of prices $\mathbf{p}$.
  - We can impose restrictions on the demand curve, such as constant slope or constant elasticity.
- With these in hand, we can:
  - Recover estimates of marginal cost (remember these are proprietary firm information and accounting estimates are often unreliable).

Overview

Multiproduct Demand system estimation is probably the most important contribution from the New Empirical IO literature

- Demand is an important primitive (think about Econ 101)
- Welfare Analysis, Pass through of taxation, Value of Advertising, Price Effects of Mergers all rely on demand estimates.
- Last 5-10 years has seen successful export to other fields: Trade, Healthcare, Education, Urban Economics, Marketing, Operations Research.
- Anywhere consumers face a number of options and "prices": doctors, hospitals, schools, mutual funds, potential dates, etc.
- There are many many more applications.

Two Major Issues

- Endogeneity of Prices
  - Prices are not randomly determined, but set strategically by firms who observe the demand curve
  - The simultaneity of supply and demand creates a problem. We see the market clearing (P^*, Q^*) over several periods, but in general we do not know which curve shifted.
- Multiple Products/Flexibility
  - We want to allow for flexible (data driven) substitution across products but if we have J products, then $\partial Q_j / \partial P_k$ might have J^2 elements.
  - We may also think that $\partial Q_j / \partial P_k$ varies with P and with other covariates x.
  - We may also care about $\partial^2 Q_j / \partial P_k^2$, $\partial^3 Q_j / \partial P_k^3$ and so on.
- We will address the issues separately and then see how to put them back together.

Taxonomy of Demand Systems

- Representative Consumer vs. Heterogeneous Agents?
- Discrete Choices vs. Continuous Choices?
- Single Product vs. Many Products?
- Product Space vs. Characteristic Space?
    - Do consumers choose products in product space or in characteristic space?

Data Sources

- We can either have aggregate data (market level data on (P_j, Q_j))
  - Many supermarket "scanner' datasets: Nielsen, IRI, Dominick's
  - NPD: video games/computers/consumer electronics.
  - Many proprietary single firm sources
- or micro data panel of data with same individuals over time.
  - Best example is Nielsen Homescan Consumer Panel data.
  - Visa/Mastercard datasets
  - Medicare
- Sometimes we have a combination of both
  - Often we have aggregate purchase data plus some micro data from a survey on a subpopulation.
  - ie: asking people who purchased GM cars which other cars they were considering.
  - Using scanner data for alcohol purchases and comparing to consumption surveys by income and education.

Economic Models of Demand

A Benchmark

Let's start with the following as a benchmark:

- A representative agent demand system.
- The consumer chooses an expenditure level for each good and consumes at least a little of all goods.
- Which desirable properties?:
  - We want a fully flexible matrix of demand derivatives $\Omega(\mathbf{p})$.
  - Probably we want some flexibility so that $\Omega(\mathbf{p}) \neq \Omega(\mathbf{p}')$.
  - Would satisfy axioms of consumer theory (WARP, Slutsky Symmetry, etc.).

Brief Aside: Constant Elasticity Demand

One candidate from your first year course would be a constant elasticity demand model. Which we could micro-found with utility for consuming $q(\omega)$ for each of J goods:

$$U = \left(\int_0^J q(\omega)^\rho d\omega\right)^{\frac{1}{\rho}} \quad 0 \leq \rho \leq 1$$

We can solve Lagrangians and find (Frisch) demands:

$$q(\omega) = \left(\frac{\lambda p(\omega)}{\rho}\right)^{\frac{1}{\rho-1}}$$

With ratios:

$$\frac{q(\omega_1)}{q(\omega_2)} = \left(\frac{p(\omega_1)}{p(\omega_2)}\right)^{\frac{1}{\rho-1}}$$

Brief Aside: Constant Elasticity Demand

Some CES algebra:

$$
\begin{aligned}
q(\omega_1) &= q(\omega_2)\left(\frac{p(\omega_1)}{p(\omega_2)}\right)^{-\sigma} \\
\underbrace{\int_0^J p(\omega_1)(\omega_1) d\,\omega_1}_{I \equiv \text{ consumer income}} &= \int_0^J q(\omega_2)p(\omega_1)^{1-\sigma}p(\omega_2)^{\sigma} d\,\omega_1 \\
I &= q(\omega_2)p(\omega_2)^{\sigma}\int_0^J p(\omega_1)^{1-\sigma} d\,\omega_1
\end{aligned}
$$

Now we can solve for Marshallian Demand:

$$
q(\omega_2) = \frac{I \cdot p(\omega_2)^{-\sigma}}{\underbrace{\int_0^J p(\omega_1)^{1-\sigma} d\,\omega_1}}
$$

Brief Aside: Constant Elasticity Demand

Using the overall price index $P = \left(\int_0^J p(\omega_1)^{1-\sigma} d\,\omega_1\right)^{\frac{1}{\rho}}$, we can re-write Marshallian demand:

$$q(\omega) \;=\; p(\omega)^{-\sigma} P^{\sigma-1} I = \left(\frac{p(\omega)}{P}\right)^{-\sigma} \frac{I}{P}$$

We can establish the well-known homotheticity property of CES by plugging back into original equation for $U(\cdot)$ and noting that $e(P, u) = P \cdot u$.

$$\begin{aligned} U &= \left(\int_0^J q(\omega)^\rho d\,\omega\right)^{1/\rho} = \left(\int_0^J p(\omega)^{1-\sigma} I^\rho P^{(\sigma-1)\rho} d\,\omega\right)^{1/\rho} \\ &= I P^{\sigma-1} \left(\int_0^J p(\omega)^{1-\sigma} d\,\omega\right)^{\frac{\sigma}{\sigma-1}} = I P^{\sigma-1} P^{-\sigma} = \frac{I}{P}. \end{aligned}$$

Brief Aside: Constant Elasticity Demand

Demand (and its derivative) for a single good:

$$
\begin{aligned}
q(p) &= p^{-\sigma} P^{\sigma-1} I \\
\frac{\partial q}{\partial p} &= -\sigma p^{-\sigma-1} P^{\sigma-1} I \\
\frac{-q}{\frac{\partial q}{\partial p}} &= \frac{p}{\sigma}
\end{aligned}
$$

So that monopoly markup becomes $p = \frac{mc}{\rho}$

- CES means one markup (and elasticity) for all goods.
- Hard to do IO here. Not so helpful in understanding strategic price setting behavior!
- Better left for Trade and Macro economists.

Almost Ideal Demand System: Deaton & Muellbauer (1980)

Recall our desirable properties:

- We want a fully flexible matrix of demand derivatives $\Omega(\mathbf{p})$.
- Probably we want some flexibility so that $\Omega(\mathbf{p}) \neq \Omega(\mathbf{p}')$.
- Would satisfy axioms of consumer theory (WARP, Slutsky Symmetry, etc.).
- Key ideas: separable preferences and multi-stage budgeting.
  - Allocating expenditures within a group: Index can be calculated without knowing what you choose within the group.
  - Other products respond only to the *index price* not to individual prices!

Begin by defining an expenditure function:

$$\log e(u,p) = (1-u)\log \underbrace{a(\mathbf{p})}_{\text{subsistence}} + u \cdot \log \underbrace{b(\mathbf{p})}_{\text{bliss}}$$

We assume a particular functional form for $a(\mathbf{p}), b(\mathbf{p})$ that is second-order flexible.

Almost Ideal Demand System: Deaton & Muellbauer (1980)

Here is the form of the expenditure function:

$$\log e(u,p) = \alpha_0 + \sum_k \alpha_k \log p_k + \frac{1}{2}\sum_k \sum_j \gamma^*_{kj} \log p_k \log p_j + u\beta_0 \prod_k p_k^{\beta_k}$$

- Estimate $(\alpha_i, \beta_i, \gamma^*_{ij})$ from data.
- We usually require $\sum_i \alpha_i = 1$, $\sum_k \gamma^*_{jk} = \sum_j \beta_j = 0$ so that demand is linearly homogenous in $\mathbf{p}$.
- Also often impose that $\gamma^*_{jk} = \gamma^*_{kj}$.
  - Sometimes we impose this *ex-ante*, other times we test for it *ex post*.
- We can also see that we have at least one parameter for each of the first two own and cross price derivatives of $e(\cdot)$.

Almost Ideal Demand System: Deaton & Muellbauer (1980)

After applying Shepard's Lemma and logarithmic differentiation, we can obtain the expenditure share for good i:

$$\begin{aligned} w_i &= \alpha_i + \sum_j \gamma_{ij} \log p_j + \beta_i u \beta_0 \prod_k p_k^{\beta_k} \quad \text{with} \quad \gamma_{ij} = \frac{1}{2}(\gamma_{ij}^* + \gamma_{ji}^*) \\ &= \alpha_i + \sum_j \gamma_{ij} \log p_j + \beta_i \log(x/P) \end{aligned}$$

- Second line comes from replacing Hicksian expenditure $h(p,u)$ with Marshallian demand.
- x represents total expenditure within group, P is the price index for the group.
- Two price indices are commonly used ("Exact" and Stone 1954's linear approximate index):

$$\begin{aligned} \log P &= \alpha_0 + \sum_k \alpha_k \log p_k + \frac{1}{2} \sum_j \sum_k \gamma_{kj} \log p_k \log p_j \\ \log P &= \sum_k w_k \log p_k \end{aligned}$$

Notes on AIDS

- AIDS seemed like a better name in 1980 than it does today!
- Gets used often in international trade or macro-consumption literature.
    - Product categories are often: durables, non-durables, housing, utilities, etc. from CEX data.
- Can use it for IO purposes (each "group" contains a single product).
- If p_k changes demand for good j (it does!) then we need an instrument for every price!
- We still have J^2 possible elasticities or $J \times (J+1)/2$.
    - Can simplify with multi-stage budgeting. (but we have to know what segments are)
    - Massive data requirements: $J = 45$ in a vending machine means we need over 2000 observations.

Beer Example: Hausman, Leonard, Zona (1994)

Goals:

- Estimate demand for beer in the US.
- Analyze a merger, test assumptions about firm conduct

Three stages:

1. Brand-Level (AIDS): 5 brands per segment.

$$\underbrace{w_i}_{\text{brand expenditure share}} = \alpha_i + \sum_j \alpha_{ij} \log p_j + \beta_i \log\left(\frac{x}{P}\right) + \varepsilon_1$$

2. Segment-Level (log-log): Premium, Light, Popular.

$$\underbrace{\log q_m}_{\text{seg. quantity}} = \beta_m \underbrace{\log y_B}_{\text{beer expenditure}} + \sum_k \sigma_k \log \underbrace{\pi_k}_{\text{segment price index}} + \alpha_m + \varepsilon_2$$

Identification: Hausman, et.al (1994)

- Price is correlated with both unobserved product quality and unobserved demand shocks.
- Finding brand level instruments is the challenge.
- The famous Hausman instrument: use prices in one city to instrument for prices in another

$$\log p_{jnt} = \delta_j \log c_{jt} + \alpha_{jn} + \omega_{jnt}$$

- Instruments tend to be strong but relevance can be questionable.
- Key is that ω_{jnt} are independent of each other (is this believable?).
    - People mostly complain about national ad campaigns (this is beer after all!)
- What about other instruments? (Input prices, taxes, etc.).
- Specification Test: brand price in other segments should not have an effect controlling for the price index of other segments.

Hausman, et.al (1994)

TABLE 1

Beer Segment Conditional Demand Equations.

	Premium	Popular	Light
Constant	0.501	−4.021	−1.183
	(0.283)	(0.560)	(0.377)
log (Beer Exp)	0.978	0.943	1.067
	(0.011)	(0.022)	(0.015)
log ($P_{PREMIUM}$)	−2.671	2.704	0.424
	(0.123)	(0.244)	(0.166)
log ($P_{POPULAR}$)	0.510	−2.707	0.747
	(0.097)	(0.193)	(0.127)
log (P_{LIGHT})	0.701	0.518	−2.424
	(0.070)	(0.140)	(0.092)
Time	−0.001	−0.000	0.002
	(0.000)	(0.001)	(0.000)
log (# of Stores)	−0.035	0.253	−0.176
	(0.016)	(0.034)	(0.023)

Number of Observations = 101.

Hausman, et.al (1994)

TABLE 2

Brand Share Equations: Premium.

	1 Budweiser	2 Molson	3 Labatts	4 Miller	5 Coors
Constant	0.393	0.377	0.230	−0.104	–
	(0.062)	(0.078)	(0.056)	(0.031)	–
Time	0.001	−0.000	0.001	0.000	–
	(0.000)	(0.000)	(0.000)	(0.000)	–
log (Y/P)	−0.004	−0.011	−0.006	0.017	–
	(0.006)	(0.007)	(0.005)	(0.003)	–
log ($P_{Budweiser}$)	−0.936	0.372	0.243	0.150	–
	(0.041)	(0.231)	(0.034)	(0.018)	–
log (P_{Molson})	0.372	−0.804	0.183	0.130	–
	(0.231)	(0.031)	(0.022)	(0.012)	–
log ($P_{Labatts}$)	0.243	0.183	−0.588	0.028	–
	(0.034)	(0.022)	(0.044)	(0.019)	–
log (P_{Miller})	0.150	0.130	0.028	−0.377	–
	(0.018)	(0.012)	(0.019)	(0.017)	–
log (# of Stores)	−0.010	0.005	−0.036	0.022	–
	(0.009)	(0.012)	(0.008)	(0.005)	–
Conditional Own	−3.527	−5.049	−4.277	−4.201	−4.641
Price Elasticity	(0.113)	(0.152)	(0.245)	(0.147)	(0.203)

$$\Sigma = \begin{Bmatrix} 0.000359 & -1.436\text{E}-05 & -0.000158 & -2.402\text{E}-05 \\ - & 0.000109 & -6.246\text{E}-05 & -1.847\text{E}-05 \\ - & - & 0.005487 & -0.000392 \\ - & - & - & 0.000492 \end{Bmatrix}$$

Note: Symmetry imposed during estimation.

Hausman, et.al (1994)

TABLE 3

Brand Share Equations: Popular Price.

	1 Old Milwaukee	2 Genesee	3 Milwaukee's Best	4 Busch	5 Piels Lager
Constant	0.287	0.225	−0.019	0.531	–
	(0.062)	(0.067)	(0.063)	(0.079)	–
Time	−0.000	−0.001	0.000	0.001	–
	(0.000)	(0.000)	(0.000)	(0.000)	–
log (Y/P)	0.014	−0.018	0.001	0.004	–
	(0.006)	(0.007)	(0.007)	(0.008)	–
log ($P_{\text{Old Milwaukee}}$)	−0.979	0.235	0.369	0.257	–
	(0.028)	(0.021)	(0.022)	(0.030)	–
log (P_{Genesee})	0.235	−0.698	0.222	0.205	–
	(0.021)	(0.029)	(0.022)	(0.030)	–
log ($P_{\text{Milwaukee's Best}}$)	0.369	0.222	−1.048	0.388	–
	(0.022)	(0.022)	(0.036)	(0.035)	–
log (P_{Busch})	0.257	0.205	0.388	−0.892	–
	(0.030)	(0.030)	(0.035)	(0.062)	–
log (# of Stores)	−0.044	0.122	−0.023	−0.091	–
	(0.010)	(0.011)	(0.010)	(0.012)	–
Conditional Own Price Elasticity	−4.789	−3.832	−5.813	−5.704	−3.956
	(0.109)	(0.120)	(0.164)	(0.329)	(0.465)

$$\Sigma = \begin{Bmatrix} 0.000603 & -0.000123 & -0.000137 & -8.289E-05 \\ - & 0.000515 & -0.000143 & -7.136E-05 \\ - & - & 0.000758 & -0.000109 \\ - & - & - & 0.00262 \end{Bmatrix}$$

Note: Symmetry imposed during estimation.

Hausman, et.al (1994)

TABLE 5

Overall Elasticities.

	Elasticity	Standard Error
Budweiser	−4.196	0.127
Molson	−5.390	0.154
Labatts	−4.592	0.247
Miller	−4.446	0.149
Coors	−4.897	0.205
Old Milwaukee	−5.277	0.118
Genesee	−4.236	0.129
Milwaukee's Best	−6.205	0.170
Busch	−6.051	0.332
Piels	−4.117	0.469
Genesee Light	−3.763	0.072
Coors Light	−4.598	0.115
Old Milwaukee Light	−6.097	0.140
Lite	−5.039	0.141
Molson Light	−5.841	0.148

Light Segment Own and Cross Elasticities.

	Genesee Light	Coors Light	Old Milwaukee Light	Lite	Molson Light
Genesee Light	−3.763	0.464	0.397	0.254	0.201
	(0.072)	(0.060)	(0.039)	(0.043)	(0.037)
Coors Light	0.569	−4.598	0.407	0.452	0.482
	(0.085)	(0.115)	(0.058)	(0.075)	(0.061)
Old Milwaukee Light	1.233	0.956	−6.097	0.841	0.565
	(0.121)	(0.132)	(0.140)	(0.112)	(0.087)
Lite	0.509	0.737	0.587	−5.039	0.577
	(0.095)	(0.122)	(0.079)	(0.141)	(0.083)
Molson Light	0.683	1.213	0.611	0.893	−5.841
	(0.124)	(0.149)	(0.093)	(0.125)	(0.148)

Hausman, et.al (1994): Results

- Relatively large own and cross price elasticities.
- Authors simulated partial merger analysis.
  - Hold prices of all non-merging parties fixed.
  - Solving for best-response of single-product.
  - How would full equilibrium analysis differ?
- Merger of Coors and Labatt's: Coors Markup 19.9% $\rightarrow$ 23.2% (small).
- Claim is that presence of other competitors constraints potential to raise prices. How? Why?

Other AIDS examples

Hausman (1997) aka *The Apple Cinnamon Cheerios War*.

- What is the value of a new good? How should we adjust CPI?
- Potentially HUGE issue. Why?
- Weekly cereal data.. 7 cities, 137 weeks. Three segments (adults, kids, family) with max 9 brands.
- Calculate $e(p_{-n}, p_n^*, u)/e(p, u)$. Find a virtual price p^* (or choke price) that leaves consumers as well off as a world without Apple-Cinnamon Cheerios.
- Virtual price is about $2\times$ actual price. CPI may be overstated by as much as 25% for all cereal brands (tons of new products).

Other AIDS examples

Chaudhuri, Goldberg, Jia (AER 2006)

- Indian market for antibiotics: (foreign vs. domestic) (licensed vs. unlicensed producers).
- Different brands, packages, etc. also different active ingredients ($J = 300$ they aggregate to four active ingredients $\times$ country of origin).
- Monthly sales data (SKU level) for 4 regions in India (Market Research firm).
- What would prices and quantities look like if intellectual property rights were enforced and unlicensed producers were shut down?

Chaudhuri, Goldberg, Jia (AER 2006)

Issues

- Products enter and exit the market. How do we model this?
- Dosages differ across products. How do we construct Q?
- Don't treat licensed v. unlicensed as different products. Why?

Results

- Estimate AIDS demand aggregated across demands
- Get upper and lower bounds on marginal costs
  - Assume that $p = mc$
  - Assume monopoly pricing.
- Calculate the virtual price or "choke price" that makes expenditures zero on unlicensed products.
- Get changes in consumer surplus (integrated demand curve) and producer profits without unlicensed firms.