Empirical Demand–Supply analysis

- Most of empirical IO focuses on estimation of demand–supply models
- Goal: say something about firm behavior (pricing, advertising, R&D, output decisions)
- While IO theory mostly concerned about supply, data constraints (lack of cost data) focus attention on demand. Fundamental tension: high prices resulting from market power ("bad") or high costs?
- On the other hand, demand characteristics (slopes) crucial to firm behavior.
- Brief review of linear simultaneous equations: basic endogeneity problem in demand analysis
- Demand analysis for differentiated product markets

Review of Simultaneous Equations

- Review linear simultaneous equations theory in specific context: estimating demand and supply
- Price endogeneity: fundamental problem in much applied IO work
- Appropriate instruments for price in supply ("demand shifters") and demand functions ("cost shifters")
- Estimation methods: IV methods (2SLS, 3SLS); Maximum likelihood

Linear Supply–demand model

- $$\begin{aligned} \text{Demand:} \quad q_t &= \gamma_1 p_t + \mathbf{x'_{t1}}\beta_\mathbf{1} + u_{t1} \\ \text{Supply:} \quad p_t &= \gamma_2 q_t + \mathbf{x'_{t2}}\beta_\mathbf{2} + u_{t2} \end{aligned}$$

- Demand function summarizes consumer preferences; supply function summarizes firms' cost structure

- If u_1 correlated with u_2, then p_t is endogenous in demand function, and q_t is endogenous in supply relation: cannot estimate using OLS.

- Graph

- x's are exogenous variables:

  1. x_{t1} are *demand shifters*; affect willingness-to-pay, but not a firm's production costs. Correlated with q_t but not with u_{2t}: use as instruments in supply function. Graph.
  2. x_{t2} are *cost shifters*; affect production costs. Correlated with p_t but not with u_{t1}: use as instruments in demand function. Graph.

Estimation methods (brief) 1

- $$\begin{aligned} \text{Demand:} \quad & q_t = \gamma_1 p_t + \mathbf{x'_{t1}}\beta_{\mathbf{1}} + u_{t1} \\ \text{Supply:} \quad & p_t = \gamma_2 q_t + \mathbf{x'_{t2}}\beta_{\mathbf{2}} + u_{t2} \end{aligned}$$

$$\Rightarrow \begin{pmatrix} 1 & -\gamma_1 \\ -\gamma_2 & 1 \end{pmatrix} \begin{pmatrix} q \\ p \end{pmatrix} = \begin{pmatrix} \beta_1 & 0 \\ 0 & \beta_2 \end{pmatrix} \begin{pmatrix} x_{t1} \\ x_{t2} \end{pmatrix} + \begin{pmatrix} u_{t1} \\ u_{t2} \end{pmatrix}$$

$$\Rightarrow Y\Gamma = XB + U$$

Estimation methods (brief) 2

- IV methods (GMM): population moment conditions $E(u_1 \cdot \mathbf{X}) = 0$, $E(u_2 \cdot \mathbf{X}) = 0$ hold at true parameter values.
  1. Two-stage least squares
  2. Three-stage least squares
- Maximum likelihood:
  1. Make distributional assumptions about $\{(u_{t1}, u_{t2})_{t=1}^{T}\}$. Example: $(u_{t1}, u_{t2}) \sim$ i.i.d $N(0, \Sigma)$
  2. Likelihood function of the data is **joint density of the endogenous variables** $(\mathbf{q_t}, \mathbf{p_t})$ **conditional on exogenous variables** $(\mathbf{x_{t1}}, \mathbf{x_{t2}})$:

$$f(Y) = g(Y\Gamma - XB) * \mid \Gamma \mid \Rightarrow$$

$$\log L(Y \mid X) \sim T \log \mid \Gamma \mid - \frac{T}{2} \log \mid \Sigma \mid$$

$$- \frac{1}{2} (Y\Gamma - XB)' \, \Sigma^{-1} \, (Y\Gamma - XB)$$

  3. Maximize this with respect to Γ, B, Σ.

Modelling demand for differentiated products

Much of recent empirical IO work: focus on analyses of differentiated product markets.

Main reasons:

1. Policy work relates to differentiated oligopolistic industries
2. Look at pricing behavior: need to model how demand would change in response to price changes.
3. Applications: autos (effects of trade restrictions), airlines (effect of hubbing, consumer heterogeneity), pharmaceuticals (effect of generics on branded prices)

- Overview: approach to modelling demand. Pitfalls thereof.
- Discrete–choice approach
- Berry/Levinsohn/Pakes methodology (price endogeneity)
- Extensions

Discrete-choice approach 1

- There are N alternatives in market. Each purchase occasion, each consumer divides her income y on (at most) one of the alternatives, and on an "outside good":

$$\max_{n,z} U(x_n, z) \text{ s.t. } p_n + p_z z = y$$

  where

  - x_n are chars of brand n, and p_n the price
  - z is quantity of outside good, and p_z its price

- Substitute in the budget constraint ($z = \frac{y-p_n}{p_z}$) to derive *conditional indirect utility functions* for each brand:

$$U_n^*(p_n, p_z, y) = U(x_n, \frac{y - p_n}{p_z}).$$

  Note: if none of the brands are bought:

$$U_0^*(p_z, y) = U(0, \frac{y - p_n}{p_z}).$$

- Consumer chooses the brand yielding the highest cond. indirect utility:

$$\max_n U_n^*(p_n, p_z, y)$$

Discrete-choice approach 2

- U_n^* usually specified as sum of deterministic and stochastic part:

$$U_n^*(p_n, p_z, y) = V_n(p_n, p_z, y) + \epsilon_n$$

Distn assumptions on ϵ_n, $n = 0 \ldots N$ determine the form that demand functions take:

$$D_n(p_1 \ldots p_N, p_z, y) = \text{Prob}\left\{\epsilon_0, \ldots, \epsilon_N : U_n^* > U_j^* \text{ for } j \neq n\right\}$$

This is also the likelihood function for an observed purchase of brand n.

Discrete-choice approach 3

Common assumptions:

- $(\epsilon_0, \ldots, \epsilon_N)$ distributed multivariate normal: **multinomial probit**. Computationally burdensome (Keane, McFadden)
- $(\epsilon_0, \ldots, \epsilon_N)$ distributed *i.i.d.* type II extreme value: multinomial logit

  $$D_n(\cdots) = \frac{\exp(V_n)}{\sum_{n'=1,\ldots,N} \exp(V_{n'})}$$

  Normalize $V_0 = 0$.

- Restrictiveness of multinomial logit: Odds ratio between any two brands i, j doesn't depend on number of alternatives available

  $$\frac{D_i}{D_j} = \frac{\exp(V_i)}{\exp(V_j)}$$

  Example: Red bus/blue bus problem. Implication: invariant to introduction (or elimination) of some alternatives. **Independence of Irrelevant Alternatives**

- Restrictive substitution patterns: $\epsilon_{i,k} = \epsilon_{j,k}$, for all $i, j \neq k$. If $V_i = \beta_i + \alpha(y - p_i)$, then $\epsilon_{i,k} = -\alpha p_k D_k$, for all $k \neq i$. Price decrease in brand k attracts proportionate chunk of demand from all other brands.

Discrete-choice approach 4

Changes to logit framework to overcome IIA:

- Nested logit: assume particular correlation structure among $(\epsilon_0, \ldots, \epsilon_N)$. Within-nest brands are "closer substitutes" than across-nest brands.
- Mixed logit: assume ϵ_n composed of two components $\epsilon_n = \theta_{1n} + \theta_{2n}$. θ_{1n}'s not necessarily *i.i.d.* across brands. Conditional on θ_{1n}'s, θ_{2n}'s are *i.i.d.* extreme value.
- Demand function is then:

$$D_n = L(n \mid \vec{p}, y) = \int L(n \mid \vec{p}, y, \vec{\theta_1}) dF(\vec{\theta_1})$$
$$= \int \frac{\exp(V_n + \theta_0)}{\sum_{n'=1,\ldots,N} \exp(V_{n'} + \theta_{n'})} dF(\vec{\theta_1})$$

  Random effects approach: assumes that $(\theta_0, \ldots, \theta_N)$ are uncorrelated with the covariates (prices) in the model.

- Berry-Levinsohn-Pakes: deal with price endogeneity

Berry-Levinsohn-Pakes methodology 1

Main idea: Control for price endogeneity in *aggregate* discrete-choice framework.

Background: Trajtenberg's study of demand for CAT scanners. Disturbing finding: coefficient on price is *positive*, implying that people prefer more expensive machines!

Explanation: quality differentials across products not adequately controlled for. In equilibrium of a diff'd product market where each product is valued on the basis of its characteristics, brands with highly-desired characteristics (higher quality) command higher prices. Unobserved quality leads to price endogeneity.

This is the type of price endogeneity tackled by Berry-Levinsohn-Pakes.

- aggregate demand model
- price endogeneity problem
- supply side
- extensions