Differentiated Products Supply and Demand

Phil Haile
Yale University

January 2021

Oligopoly Supply

Baseline workhorse model:

- firms produce differentiated goods/products, selling to consumers with heterogeneous preferences
- static model, complete information (model of long run eqm)
  - set of products, their non-price characteristics already set
  - Nash eqm in simultaneous price setting game in each market
    - market usually defined by time or geography
    - can be further restricted to a set of consumers (e.g., college-educ female in CT in 2020, age$<$30, income$<$ \$50K).

Firm Cost Functions

variable cost $C_j(Q_j, \mathrm{w}_{jt}, \omega_{jt}, \gamma)$ for product j

- Q_j total quantity of good j sold
- w_{jt} observable cost shifters; may include product characteristics x_{jt} that will affect demand (later)
- ω_{jt} unobserved cost shifters ("cost shocks"); may be correlated with latent demand shocks (later)
- γ parameters
- for multi-product firms, we'll assume variable cost additive across products for simplicity

We ignore fixed costs: these affect entry/exit/innovation (later this week) but not pricing conditional on these things.

Demand

Notation

- J_t products/goods/choices in market t (for now $J_t = J$)
- $p_t = (p_{1t}, \ldots, p_{Jt})$, prices of all goods
- $\chi_t = (\chi_{1t}, \ldots, \chi_{Jt})$, other characteristics of goods affecting demand (observed and unobserved to us)

Demand system:

$$q_{jt} = q_j\left(p_t, \chi_t\right) \quad j = 1, \ldots, J.$$

Equilibrium Pricing

With single-product firms and constant marginal cost (simple case)

$$\pi_{jt} = q_j\left(p_t, \chi_t\right)\left[p_{jt} - mc_j\left(\mathrm{w}_{jt}, \omega_{jt}, \gamma\right)\right]$$

FOC wrt to p_{jt} :

$$p_{jt} = mc_{jt} - q_j\left(p_t, \chi_t\right)\left(\frac{\partial q_j}{\partial p_{jt}}\right)^{-1}$$

This is *inverse elasticity pricing* (i.e., monopoly pricing) against the "residual demand curve" $q_j\left(p_t, \chi_t\right)$:

$$\frac{p_{jt} - mc_{jt}}{p_{jt}} = -\frac{q_j\left(p_t, \chi_t\right)}{p_{jt}}\left(\frac{\partial q_j}{\partial p_{jt}}\right)^{-1}.$$

Equilibrium Pricing with Multi-Product Firms

With multi-product firms, firm f's profit is

$$\begin{aligned}\Pi_{ft} &= \sum_{j \in J_f} \pi_{jt} \\ &= \sum_{j \in J_f} q_j\left(p_t, \chi_t\right)\left[p_{jt} - mc_j\left(\mathrm{w}_{jt}, \omega_{jt}, \gamma\right)\right]\end{aligned}$$

FOC wrt p_j :

$$p_{jt} = mc_{jt} - \left(\frac{\partial q_j}{\partial p_{jt}}\right)^{-1}\left[q_j\left(p_t, \chi_t\right) + \sum_{k \in J_f \setminus \{j\}} \frac{\partial q_k}{\partial p_{jt}}\left(p_{kt} - mc_{kt}\right)\right]$$

(firm internalizes effects of Δp_j on profit from all of its products).

Supply Model

What we get from this...

1. Holding all else fixed, markups/prices depend on the own-price elasticities of residual demand. Equilibrium depends, further, on how a change in price of one good affects the quantities sold of others, i.e., on cross-price demand elasticities.

$\Longrightarrow$ For good quantitative predictions of firm behavior and market outcomes, we will need good estimates of demand (own and cross-price derivatives)

Supply Model

What we get from this...

2. If we known demand, we can also perform a small miracle:
   - re-arrange FOC above:

$$mc_{jt} = p_{jt} + q_j\left(p_t, \chi_t\right)\left(\frac{\partial q_j}{\partial p_{jt}}\right)^{-1}$$

     supply model + estimated demand→estimates of marg costs!
   - with multiproduct firms, same thing in system of equations:

$$mc_{jt} = p_{jt} + \left(\frac{\partial q_j}{\partial p_{jt}}\right)^{-1}\left[q_j\left(p_t, \chi_t\right) + \sum_{k \in J_f \backslash \{j\}} \frac{\partial q_k}{\partial p_{jt}}\left(p_{kt} - mc_{kt}\right)\right]$$

[see Rosse (1970), Bresnahan (1981, 1987), BLP (1995), Berry-Haile (2014)].

Supply Model

What we get from this...

3. If we know demand and marginal costs, we can "predict" a lot of stuff—i.e., give the quantitative implications of the model for counterfactual worlds: e.g., what prices, consumer choices, profits, consumer welfare ... if
   - a tax or tariff were imposed?
   - two suppliers merged?
   - a certain new product had not been introduced?
   - school vouchers were provided to poor students?

Demand can be important on its own. But good demand estimates open a world of possibilities for answering questions about markets and competition policy.

Demand Isn't Easy

Typically we need to know levels/elasticities of demand at particular points; i.e., effects of one price change holding all else fixed

The main challenge: unobserved demand shifters ("demand shocks") at the level of the good×market (e.g., unobserved product char or market-specific variation in mean tastes for products)

- demand shocks are among the things that must be held fixed to measure the relevant demand elasticities etc.
- explicit modeling of these demand shocks central in the applied IO literature following Berry-Levinsohn-Pakes 1995 (often ignored outside this literature).

A Key Challenge

Econ 101: the quantity demanded of a given good j depends on the prices and characteristics of all related goods (substitutes and complements). This includes the latent demand shocks associated with all of those goods.

So with J related goods, demand for each one takes the form

$$q_j = D(x, p, \xi)$$

where

- p is a J-vector of all goods' prices.
- x is the matrix of all non-price observables
- ξ is a J-vector of demand shocks for all goods.

Demand Is Not Regression

$$q_j = D_j(x, p, \xi) \tag{1}$$

- RHS has many endogenous variables and many latent shocks;
- D_j is a standard regression function only under a strong functional form assumption: that the J components of ξ enter D_j only through a scalar index
- not obvious how to proceed, even if prices were exogenous!
- applying regression methods to (1) might allow one to recover certain weighted average derivatives of demand, but those have little, if any, value

$\Longrightarrow$ **We have to approach demand differently.**

Price Endogeneity Adds to the Challenge

- all J endogenous prices are on RHS of demand for each good
- eqm pricing implies that each price depends on all demand shocks and all cost shocks
  - $\rightarrow$ prices endogenous
  - $\rightarrow$ control function generally is not a valid solution
- clear that we need sources of exogenous price variation, but
  - what exactly is required?
  - how do we proceed?

Ways Forward

By far the most common approach starts by building up a demand system from a smaller set of parameters appearing in a specification of consumer utilities.

Deriving demand from utilities offers parsimony: many own- and cross-price elasticities from a modest number of parameters. And we'll use a "trick" (a useful mathematical result) to deal with the fact that each good's demand is affected by J structural errors.

Specifying utilities is not essential (although, typically, very useful); the "trick" (or some other trick) is.

NEXT: RANDOM UTILITY DISCRETE CHOICE

Discrete Choice Demand

- consumers have unit demands
- each chooses one of the available options
- generally, one option should be "none of the above" — what we will call the "outside good" (without this, there would be no aggregate demand elasticity!)

Note: discrete choice is more general than it seems; e.g., a single option for a consumer could be $\{$Dodge Caravan + Porsche 911$\}$ or $\{$four boxes of cookies + a gallon of milk$\}$. However, key insights we'll cover here extend to models of "multiple discrete choice" or continuous choice.

Random Utility Discrete Choice Demand

Random Utility Specification

- differentiated goods $j \in \{1, \ldots J\}$
- *conditional indirect utilities* of consumer i: u_{ij} ("utility")
- $(u_{i1}, \ldots, u_{iJ}) \sim F_u(\cdot)$ for all i
- outside good 0
  - only utility differences matter, so we could set $u_{i0} = 0 \forall i$ (or give it any distribution we want) wlog.

Choice Probabilities From the Model

- consumer i's choice (quantities)

$$q_{ij} = 1\{u_{ij} \geq u_{ik} \ \forall k = 0, \ldots, J\}$$

(typical assumptions imply ties happen w/probability zero)

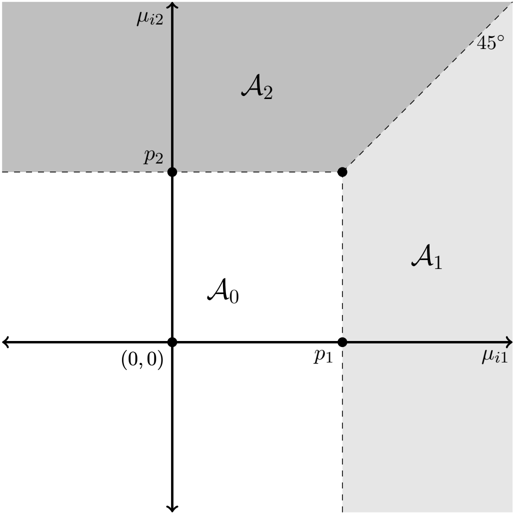
- yields choice probabilities

$$\begin{aligned} s_{ij} &= \Pr\left(q_{ij} = 1\right) \\ &= \int_{\mathcal{A}_j} dF_U\left(u_{i1}, \ldots, u_{iJ}\right) \end{aligned}$$

where

$$\mathcal{A}_j = \left\{ \left(u_{i1}, \ldots, u_{iJ}\right) \in \mathbb{R}^J : u_{ij} \geq u_{ik} \ \forall k \right\}.$$

Example: $J = 2, u_{ij} = \mu_{ij} - p_j$



Demand and Utility

- utility maximization is a convenient way to represent/rationalize demand (consumer choice rules)
- but utility is a notion we make up; to define/estimate demand, one need even not assume the conditions that permit utility maximization to rationalize behavior
- indeed, randomness in the "utilities" could reflect noise/errors in consumer choice (e.g., Luce, 1959)
    - profit-maximizing firms don't care what the randomness represents (unless they can affect it)
    - but the interpretation will matter for welfare.

NEXT: BLP DEMAND MODEL

Berry, Levinsohn, and Pakes (1995) "BLP"

Standard empirical model of demand and supply of differentiated products.

Many of the ideas also in Berry (1994), mostly for simpler models.

Many extensions and variations.

BLP Random Utility Specification

(slightly simplified)

$$u_{ijt} = x_{jt}\beta_{it} - \alpha p_{jt} + \xi_{jt} + \epsilon_{ijt}$$

- consumer i, good/product j, market t
  (best to imagine data from many markets, each with many consumers)
- $x_{jt} \in \mathbb{R}^K, p_{jt}$ *observable* product/market characteristics
- ξ_{jt} *unobserved* demand shock at level of product×market
- ϵ_{ijt} idiosyncratic (and latent) "taste for product"

...

Preference Heterogeneity

$$u_{ijt} = x_{jt}\beta_{it} - \alpha p_{jt} + \xi_{jt} + \epsilon_{ijt}$$

Sources of csr heterogeneity: $\epsilon_{ijt}, \beta_{it} = \left(\beta_{it}^{1}, \ldots, \beta_{it}^{K}\right)$

- $\beta_{it}^{k} = \beta_{0}^{k} + \sigma_k \zeta_{it}^{k}$ (“random coefficient” $=$ taste for $x_{jt}^{(k)}$)
- $\left\{\epsilon_{ijt}, \zeta_{it}^{k}\right\}_{j,k}$, i.i.d. across csrs and mkts
- typically:
  - $\epsilon_{ijt} \sim$ i.i.d. type 1 extreme value (like multinomial logit)
  - $\zeta_{it}^{k} \sim$ i.i.d. standard normal, or drawn from actual distribution of demographics (e.g., income) in market t.

Exogenous and Endogenous Product Characteristics

Recall

$$u_{ijt} = x_{jt}\beta_{it} - \alpha p_{jt} + \xi_{jt} + \epsilon_{ijt} \tag{2}$$

- exogenous characteristics: $x_{jt} \perp\!\!\!\perp \xi_{jt}$
- endogenous characteristics: p_{jt} (usually a scalar, price)
  - typically each p_{jt} will depend on *whole vector* $\xi_t = (\xi_{1t}, \ldots, \xi_{Jt})$ (and on own and others' costs)
  - we need to distinguish true effects of prices on demand from the effects of ξ_t; this will require instruments
  - of course (2) is not an estimating equation (u_{ijt} not observed)
  - because prices and quantities are all endogenous—indeed, determined—*simultaneously*, you may suspect (correctly) that instruments for prices alone may not suffice.

Utility Specification, Rewritten

Rewrite

$$\begin{aligned} u_{ijt} &= x_{jt}\beta_{it} - \alpha p_{jt} + \xi_{jt} + \epsilon_{ijt} \\ &= \delta_{jt} + \nu_{ijt} \end{aligned}$$

where

$$\begin{aligned} \delta_{jt} &= x_{jt}\beta_0 - \alpha p_{jt} + \xi_{jt} \quad (\text{“mean utility” of good } j \text{ in market } t) \\ \nu_{ijt} &= \sum_k x_{jt}^k \sigma^k \zeta_{it}^k + \epsilon_{ijt} \\ &\equiv x_{jt}\tilde{\beta}_{it} + \epsilon_{ijt} \quad (\text{defining } \tilde{\beta}_{it}\text{—the random part of } \beta_{it}). \end{aligned}$$

Market Shares

- recall $u_{ijt} = \delta_{jt} + \nu_{ijt}$
- let $\delta_t = (\delta_{1t}, \ldots, \delta_{Jt})$, $\delta_{0t} = 0$ (the normalization mentioned)
- $\approx$ continuum of consumers in each market*
  $\Longrightarrow$ market shares = choice probabilities =

$$s_{jt} = \Pr(y_{it} = j) = \int_{\mathcal{A}_j(\delta_t)} dF_\nu(\nu_{i0t}, \nu_{i1t}, \ldots, \nu_{iJt})$$

where

$$\mathcal{A}_j(\delta_t) = \left\{ (\nu_{i0t}, \nu_{i1t}, \ldots, \nu_{iJt}) \in \mathbb{R}^{J+1} : \delta_{jt} + \nu_{ijt} \geq \delta_{kt} + \nu_{ikt} \ \forall k \right\}$$

* really, enough that sampling error on choice probs negligible compared to that of moments based on variation across products/markets.

Demand

- market shares again

$$s_{jt} = \int_{\mathcal{A}_j(\delta_t)} dF_\nu \left(\nu_{i0t}, \nu_{i1t}, \ldots, \nu_{iJt}\right)$$

- with random coefficients, $F_\nu(\cdot)$ is really $F_\nu(\cdot | x_t, \sigma)$ where
  - $x_t = (x_{1t}, \ldots, x_{Jt}) \in \mathbb{R}^{K \times J}$
  - $\sigma = (\sigma^1, \ldots, \sigma^K)$
- so $s_{jt} = s_j(\delta_t, x_t, \sigma)$
- if M_t is the total measure of consumers in market t, quantities demanded are

$$q_{jt} = M_t \times s_j(\delta_t, x_t, \sigma).$$

Discussion

Key features of the BLP model

- explicit modeling of demand shocks
- consumer heterogeneity through random coefficients

We discussed the need to be explicit about demand shocks. Why random coefficients?

Why Random Coefficients?

Without random coefficients:

$$
\begin{aligned}
u_{ijt} &= \underbrace{x_{jt}\beta_0 - \alpha p_{jt} + \xi_{jt}}_{} + \epsilon_{ijt} \\
&= \qquad \delta_{jt} \qquad + \epsilon_{ijt}
\end{aligned}
$$

If ϵ_{ijt} are iid and independent of (x, p), e.g. as in the multinomial logit model, products differ only in δ_{jt}

- $\implies$ market shares depend only on the mean utilities;
- $\implies$ price elasticities (own and cross) depend only on mean utilities too.

Example

Two autos with (virtually) identical market shares in 2020:

	MSRP (base)	Mkt Share
Automobile 1		0.01%
Automobile 2		0.01%

Without random coefficients, model implies same mean utility for each and therefore: same own-price demand elasticity, same cross-price elasticity wrt price of any third automobile, say Ford F-150 pickup (#1 market share) or Toyota Camry (best-selling car).

Example

Two autos with (virtually) identical market shares in 2020:

	MSRP (base)	Mkt Share
Toyota Corolla	$20,000	0.01%
GMC Sierra Pickup	$30,000	0.01%

Without random coefficients, model implies same mean utility for each and therefore: same own-price demand elasticity, same cross-price elasticity wrt price of any third automobile, say Ford F-150 pickup (#1 market share) or Toyota Camry (best-selling car).

How do random coefficients help?

Real goods differ in multiple dimensions; real consumers have (heterogeneous) preferences over these differences

- random coefficients on product characteristics can capture this
  - large $\beta_i^k \iff$ strong taste for characteristic x^k (e.g., fuel efficiency or dummy for pickup)
  - i's first choice likely to have high value of x^k
  - i's second choice too!
    (note: cross elasticities are always about 1st vs. 2nd choices)
- incorporating this allows more sensible substitution patterns: competition is mostly "local" – i.e., between firms offering products appealing to the same consumers.

Which random coefficients?

We must choose which characteristics have random coefficients

- dummies for subsets of products?
    - this covers the *nested logit* as a special case: see Berry (1994)
- certain horizontal or vertical characteristics (parts of X, P)?

In practice, the choice depends on the application and data set, including instruments. Too many RC's for the data available will often yield imprecise estimates of σ.

NEXT: ESTIMATION OF THE BLP DEMAND MODEL

Estimation with Market-Level Data: A Partial Sketch

Observables: x_t, p_t, sh_t,w_t, and $\tilde{z}_t$ $\leftarrow$excluded IV
(for clarity, sh_{jt} and sh_t denote the *observed* market shares)

1. start with demand model alone
2. suppose $F_\nu(\cdot|x_t, \sigma)$ is known (i.e., σ known)
3. for each market t, find $\delta_t \in \mathbb{R}^J$ such that $s_j(\delta_t, x_t, \sigma) = sh_{jt}\ \forall j$
   i.e., "invert" model at observed market shares to find mean utilities δ_t
4. using IV ($E[Z_{jt}|\xi_{jt}] = 0$), estimate the equation

$$\delta_{jt} = x_{jt}\beta_0 - \alpha p_{jt} + \xi_{jt}.$$

Some Details to Fill In

ok . . . a lot of details

1. What instruments?
2. Will the "inversion" step actually work?
3. What about σ ??
4. Formally define estimator and computational algorithm(s)
5. Add Supply Side
    - additional restrictions (moment conditions) aid estimation of demand
    - estimate parameters γ of marginal cost function (why? may care directly; and needed for counterfactuals that change equilibrium quantities unless *mc* is constant).

NEXT: INSTRUMENTS

Instruments for Estimating Demand

Broadly speaking, we need variables that exogenously shift all endogenous variables—*prices and quantities*—*independently*.

This may be counterintuitive: to estimate demand, we might think instruments for prices were all we needed. As we discussed earlier, however, exogenous variation in prices generally doesn't suffice. More below.

Typical (Excluded) Instruments for Estimating Demand

1. Excluded cost shifters w_t
   - classic demand instrument, e.g., wages, material costs, shipping cost to market t, taxes/tariffs, demand shifters from other markets
2. Proxies for excluded costs shifters
   - typical: price of same good in another mkt ("Hausman instruments"); properly excluded if demand shocks in one market not correlated with those in others

Typical (Excluded) Instruments for Estimating Demand

3. Markup shifters: e.g., characteristics of "nearby" markets ("Waldfogel instruments")
   - e.g., firms may use same price for all markets in a region
   - e.g., age/income/education in San Francisco may affect prices (markups) in Oakland, but may be independent of Oakland preferences (including Oakland demand shocks) conditional on Oakland observables
4. "BLP Instruments" x_{-jt}
   - by assumption, $E\left[\xi_{jt}|x_t\right] = E\left[\xi_{jt}\right]$
   - affect quantities directly; affect prices (markups) via eqm

Later: optimal functions of the excluded instruments for the unconditional moments.

NEXT: INVERSION

Will the Inversion Step Work?

Given x, σ and any positive shares sh, define $\Phi : \mathbb{R}^J \to \mathbb{R}^J$ by

$$\Phi(\delta) = \delta + \ln(sh) - \ln(s(\delta, x, \sigma))$$

Berry (1994) shows (under mild conditions on the linear random coefficients random utility model—extreme value and normal random coeff not necessary) that for any nonzero shares sh, Φ is a contraction, i.e.,

- it has a unique fixed point in δ
  - $s(\delta_t, x_t, \sigma)$ has an inverse: we can write $\delta_t = \delta(s_t; x_t, \sigma)$
- $\exists$ convergent algorithm: start with guess δ^0, set $c = 1$
  1. let $\delta^c = \Phi\left(\delta^{c-1}\right)$, $c = 1, 2, \ldots,$
  2. repeat to convergence.

NEXT: Identification?

What About sigma??

- inversion result $\implies$ for <u>any</u> market shares s_t and <u>any</u> σ, we can find a vector δ_t that rationalizes the data with the BLP model
- a non-identification result? there is NO information about σ from market shares?

What are we forgetting?

- cross-market (and cross-product) variation
  - we have a model of the mean utilities δ_{jt}
  - the structural errors $\xi_{jt} = \delta_{jt} - x_{jt}\beta_0 - \alpha p_{jt}$ implied by candidate (α, β, σ) and inversion must be mean-independent of exogenous observables across markets and products
- (this is just like linear regression: for any (x, y, β) $\exists e$ such that $y = x\beta + e$, but $x \perp e$ is what ensures identification of β).

Identification of sigma: loose nonparametric intuition

Changes in choice sets (in the case of cars)

- recall $s_{jt} = s\left(\delta_t, x_t, \sigma\right)$
- consider two markets, same p_t, x_t, ξ_t in each
- remove 1 car in one of them... where does its mkt share "go"?
  - to cars with large mkt shares?
  - to cars similar to the one removed in some dimension(s)?
- similar idea with cts variation across/within markets

One source of the looseness: fixing $\xi_t = (\xi_{1t}, \ldots, \xi_{Jt})$ in two markets is typically not possible. Instruments will shift things independently of ξ_t, but this isn't the same a fixing them.

Identification of sigma: parametric intuition

Counting parameters and moments:

- trial value of $\sigma \implies \delta_t(sh_t; x, \sigma)$ by inversion
- with trial value of $\alpha, \beta \implies \xi_t(\sigma, \alpha, \beta)$
- IV orthogonality condition: $E[\xi_{jt}(\sigma, \alpha, \beta) Z_{jt}] = 0 \; \forall j, t$
- what kind of Z do we need? we need at least as many moment conditions as parameters
  - $\implies x_{jt}$ plus excluded IV for each price is not enough*: (α, β) are not all the parameters!*
  - we need excluded instruments "to identify σ" too

Clear that we will need at least as many exogenous variables (instruments) as we have parameters in the model.

Nonparametric Identification of Demand

Even if considerations lead to reliance on parsimonious parametric specifications, we'd like to know what does (or does not) permit identification without such restrictions.

Berry and Haile (2014) examine a nonparametric generalization of the BLP demand model and show identification with market-level data. The main requirement: instruments creating independent exogenous variation in all $2J$ endogenous variables: prices and quantities:

- intuitively: move 1 price, holding fixed $J-1$ other prices and J demand shocks $\rightarrow 2J$
- instruments: "BLP instruments" (exogenous characteristics of other products) plus J others (shifters of costs or markups).

NEXT: ESTIMATION DETAIL

Basic Idea for BLP Estimator (demand alone)

idea: method of moments estimator

- any guess at the parameters (σ, α, β) implies values $\xi_{jt}(\sigma, \alpha, \beta)$ for the latent demand shocks rationalizing the data
- moments $\implies$ $E[\xi_{jt}(\sigma, \alpha, \beta) z_{jt}] = 0$
- sample analog $E[\xi_{jt}(\sigma, \alpha, \beta) z_{jt}] \approx \frac{1}{JT} \xi_{jt}(\sigma, \alpha, \beta) z_{jt}$
- GMM estimator
  - $\left(\hat{\sigma}, \hat{\alpha}, \hat{\beta}\right)$ chosen to make sample analog close to zero
  - optimal weighting of moments for efficiency.

Some Complications

1. model predictions $s_j(\delta_t, x_t, \sigma)$ involve high-dimensional integrals (recall 2-D picture)
   - use simulation to approximate
   - $\Longrightarrow$ "method of simulated moments"
2. moment conditions involve $\xi_t(\sigma, \alpha, \beta)$, which has no closed form $\rightarrow$two options for computation of the estimator:
   - solve contraction at each trial value of (σ, α, β) $\Longrightarrow$ "nested fixed point" algorithm (BLP)
   - forget about contraction, solve the BLP constrained optimization problem directly using specialized algorithms adapted to the BLP details (Dube-Fox-Su, 2012).

Defining the Estimator

Notation

- let $\theta = (\theta_1, \theta_2) = ([\alpha, \beta_0], \sigma)$
- let Z_{jt} denote the exogenous variables $(x_{jt}, \mathrm{w}_{jt}, \tilde{z}_{jt})$
- let $\delta_{jt}(\theta_2)$ be shorthand for $\delta_j(sh_t; x_t, \sigma)$.

The BLP Estimator

GMM estimator of θ defined as solution to mathematical program:

$$
\begin{aligned}
\min_{\theta} \quad & g(\xi(\theta))' W g(\xi(\theta)) \quad \text{s.t.} \\
g(\xi(\theta)) &= \frac{1}{N} \sum_{\forall j,t} \xi_{jt}(\theta) z_{jt} \\
\xi_{jt}(\theta) &= \delta_{jt}(\theta_2) - x_{jt}\beta - \alpha p_{jt} \\
\log(sh_{jt}) &= \log(s_j(\delta_t, x_t, \theta_2)) \\
s_j(\delta_t, x_t, \theta_2) &= \int \frac{\exp[\delta_{jt}(\theta_2) + x_{jt}\tilde{\beta}]}{1 + \sum_k \exp[\delta_{jt}(\theta_2) + x_{kt}\tilde{\beta}]} f_{\tilde{\beta}}(\tilde{\beta}|\theta_2) d\tilde{\beta}_i \\
W &= \text{standard GMM weight matrix.}
\end{aligned}
$$

NEXT: COMPUTATION OF THE BLP ESTIMATOR

BLP Estimation Algorithm (Sketch)

"Nested Fixed Point" algorithm (used for other things too)

- Outer Loop: search over trial values of θ
- Inner Loop: given θ, find solution for $\xi(\theta)$
    - given θ_2, solve for $\delta(\theta_2)$ as fixed point of contraction mapping
    - then $\xi_{jt}(\theta) = \delta(\theta_2) + \alpha p_{jt} - x_{jt}\beta_0$

```
begin outer loop
  try new θ
  begin inner loop
    solve contraction
  end inner loop
  calculate GMM criterion
end outer loop
```

"MPEC" Algorithm (Sketch)

(mathematical programming with equilibrium constraints)

Dube, Fox, and Su (2012)

- the general idea:
  - BLP estimator is defined by a constrained optimization problem: *minimize GMM objective function over parameters, subject to constraint that the inner loop fixed point equations hold*
  - so try off-the-shelf constrained optimization solvers that work well for "sufficiently nice" problems
- DFS: highlight critical details and tricks that can make the BLP-MPEC problem "sufficiently nice"
- code posted by authors.

Optimization: NFP vs. MPEC?

Naive implementation of either approach can easily fail. But both can work well when one follows now-established best practices. Both have publicly available implementations.

The open source pyBLP implementation—discussed in some detail in Conlon and Gortmaker (2020)—offers a frontier NFP approach incorporating multiple advances and options. They also have useful online tutorials (google pyBLP) that complement the published paper. I recommend starting here.

NEXT: "OPTIMAL INSTRUMENTS"

Optimal Instruments: An Important Digression

Loosely

- many possible functions of the exogenous variables X, Z could serve as instruments
- what is the best choice?
- particularly relevant to BLP IV: many subsets/combinations/functions of huge x_{-jt} could be used

Formally: which unconditional moment conditions yield asymptotic efficiency?

Optimal Instruments

- recall $\theta = (\alpha, \beta, \sigma)$; from Chamberlain (1986) the optimal (but infeasible) demand-side instruments are

$$D_{jt}(z_t) = E\left[\left.\frac{\partial \xi_{jt}(\theta^0)}{\partial \theta}\right| z_t\right]$$

- feasible approximations:
  - initially explored in BLP 1995; much improved in BLP 1999
  - simpler version in Reynaert-Verboven (2014)
  - related ideas in Gandhi-Houde (2019)
  - in practice: approximate optimal IV often help substantially
  - Conlon-Gortmaker (2020): more detail, more options, all available in pyBLP.

NEXT: BRINGING SUPPLY BACK TO THE MODEL

Adding the Supply Side Moments

- suppose

$$mc_{jt}\left(\mathrm{w}_{jt},\omega_{jt},\gamma\right)=\mathrm{w}_{jt}\gamma+\omega_{jt}$$

- recall firm FOC:

$$p_{jt}-\mathrm{w}_{jt}\gamma+\omega_{jt}-\frac{s_j\left(\delta_t,x_t,\sigma\right)}{\alpha}\left(\frac{\partial s_j}{\partial\delta_{jt}}\right)^{-1}=0$$

- so for any $(\sigma,\alpha,\beta,\gamma)$, we have an implied ω_{jt}
- additional moments for estimation:

$$E\left[\omega_{jt}\left(\sigma,\alpha,\beta,\gamma\right)\tilde{z}_{jt}\right]=0$$

- Note: supply moments depend on demand parameters too; in practice, these often help precision of demand estimates—σ in particular.

Some Extensions and Active Topics

1. multiple endogenous product characteristics (e.g., Fan, 2013)
2. incomplete "consideration sets" (e.g., Goeree, 2008)
3. "multiple discrete choice" (e.g., Hendel, 1999)
4. LASSO selection of covariates/instruments (e.g., Gillen, Montero, Moon & Shum, 2015)
5. EL estimator (e.g., Conlon, 2013)
6. nonparametric estimation/inference (e.g., Compiani, 2019)
7. discriminating between models of supply (Berry & Haile, 2014; Backus, Conlon, & Sinkinson, 2020)