Structural Estimation of Differentiated-Product Industries: Introduction of the BLP Framework

Ching-I Huang
National Taiwan University

June 23, 2014

presented at Academia Sinica

Structural Estimation v.s. Reduced-Form Approach

- Structural Approach:
  - Models are derived from economic theories.
  - Use data to estimate parameters in the model.
  - Possible to perform counterfactual experiments
  - Possible to calculate consumer's surplus and social welfare
- Reduced-Form Approach:
  - The main goal is to find a causal relationship between x and y.
  - Do not require knowledge of the underlying data generating process (i.e. economic rules characterizing agents' behavior)

What Do We Want to Know?

- Anti-trust perspective:
  - Do firms abuse their market power?
  - Should we allow a proposed merger?
  - The effects of price discrimination
  - . . .

- Marketing perspective:
  - What is the optimal pricing scheme?
  - How to introduce a new product?
  - . . .

Goal of the Lecture

- Introduce the method developed by Berry, Levinsohn, and Pakes (BLP) in “Automobile Prices in Market Equilibrium” (*Econometrica, 1995*).

- A method to empirically analyze markets with differentiated products
- A structure approach, allowing counterfactual simulations

- Recent reference: Ackerberg, Benkard, Berry, and Pakes (2007) “Econometric Tools for Analyzing Market Outcomes” in *Handbook of Econometrics*, Vol. 6A, Chap. 63, Sect. 1.

Outline of the Talk

- **Background**
- Basic Version of the BLP Framework
- Extensions of the BLP Framework
- Applications
- Technical Issues and Recent Developments
- Using the BLP Framework in Estimation

Background: Previous Approaches

- Structure-Conduct-Performance paradigm: Analyzing the relationship between the market structure and the market performance.

- Neoclassical Demand Systems: A flexible functional form to estimate the demand for differentiated products.

Structure-Conduct-Performance Paradigm

- An industry's performance depends on the conduct of sellers and buyers, which depends on the structure of the market.
- A typical regression is

$$\pi = f\left(\text{concentration}, \text{minimal efficient scale}, \frac{AD}{sales}, \frac{R\&D}{sales}, \ldots\right).$$

  - π: market performance, eg. rate of return, price-cost margin, Tobin's q (market value/asset value)
  - Measures of market structure:
    - concentration, eg. C4, HHI
    - barriers to entry
    - unionization

Structure-Conduct-Performance Paradigm: Drawbacks

- Bad data: using accounting data to measure marginal costs
- Market definition: using Standard Industrial Classification
- Regressors are endogenous.
- Demsetz's critique: More efficient firms earn more profits and also have higher market shares.

Neoclassical Demand Systems

- Almost Ideal Demand System (AIDS) for the demand of differentiated goods

$$s_i = \alpha_i + \sum_j \gamma_{ij} \log p_j + \beta_i \log(X/P),$$

  - s_i budget share for product i
  - p_i price of product i
  - X total budget
  - P a price index
- Imposing restrictions based on theory:

$$\sum_j \gamma_{ij} = \sum_i \gamma_{ij} = \sum_i \beta_i = 0, \quad \sum_i \alpha_i = 1, \quad \gamma_{ij} = \gamma_{ji}$$

Neoclassical Demand Systems: Drawbacks

- Too many parameters to be estimated
- Prices are endogenous.
- New-product problem

Outline of the Talk

- Background
- **Basic Version of the BLP Framework**
- Extensions of the BLP Framework
- Applications
- Technical Issues and Recent Developments
- Using the BLP Framework in Estimation

Main Features of the BLP Framework

- Products are bundles of characteristics.
- Preferences are defined on those characteristics.
- Each consumer chooses a bundle that maximizes its utility.
- Consumers have different preferences for different characteristics, and hence make different choices.
- Simulations is used to obtain aggregated demand.

The Basic BLP Framework: Product Characteristics

- Typically products are differentiated in many ways.
- We could not expect to obtain precise estimates of all the relevant characteristics.
- One solution is to put in the “important” differentiating characteristics, say x, and an unobservable, say ξ, which picks up the aggregate effect of the multitude of characteristics that are being omitted (by the econometrician).
- To the extent that producers know ξ when they set prices, goods that have high values for ξ will be priced higher in any reasonable notion of equilibrium.

The Basic BLP Framework: A Consumer's Utility

- Consumer i's utility of consuming product $j \in \{1, 2, \ldots, N\}$ can be expressed as

$$u_{ij} = x_j \beta_i - \alpha_i p_j + \xi_j + \varepsilon_{ij},$$

  - x_j: vector of observable characteristics of product j
  - ξ_j: unobservable characteristics of product j
  - p_j: price of product j
  - α_i, β_i, ε_{ij}: consumer-specific taste parameters.

- We also assume the existence of an outside good, $j = 0$.
  - Consumers may choose to purchase the outside good instead of one of the N "inside" products.
  - In the absence of an outside good, consumers are forced to choose from the inside good.

The Basic BLP Framework: A Consumer's Choice

- Each consumer purchases one unit of the good that gives the highest utility.
- Conditional on the characteristics (x, ξ) and prices p, a consumer chooses product j if and only if

$$u_{ij} \geq u_{ik}$$

for all $k \in \{0, 1, 2, \ldots, N\}$.

Logit Model as a Special Case

- Assume the coefficients are constant across consumers: $\alpha_i = \alpha$ and $\beta_i = \beta$.

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

- Define the average utility level of product j as

$$\delta_j \equiv x_j\beta - \alpha p_j + \xi_j.$$

- Normalize the mean utility of the outside good to zero: $\delta_0 = 0$.
- Assume ε_{ij} is type I extreme value distribution i.i.d. across i and j:

$$F(\varepsilon) = \exp(-\exp(-\varepsilon)).$$

Market Shares under Logit Model

- By Law of Large Numbers, the market share of product j is

$$s_j(\delta) = \Pr(\delta_j + \varepsilon_{ij} \geq \delta_k + \varepsilon_{ik}, \forall k \neq j) = \frac{e^{\delta_j}}{\sum_{k=0}^{N} e^{\delta_k}}.$$

- To estimate the coefficients, α and β, we could potentially run a nonlinear regression on the observed market shares using the above equation.
  - difficult to solve endogeneity problem under a nonlinear model ...
- "Invert" the market shares s_j into the implied mean utility levels δ_j:

$$\log(s_j) - \log(s_0) = \delta_j \equiv x_j\beta - \alpha p_j + \xi_j.$$

- α and β can be estimated by IV regression.

Drawbacks of the Logit Model

- The assumptions on taste heterogeneity imposes strong restrictions on the pattern of cross-price elasticities.
- Cross-price elasticities can only depend on the value of δ_j, with no additional effect from individual product characteristics or prices.

$$\frac{\partial s_k}{\partial p_l} = \alpha s_k s_l \text{ for } k \neq l.$$

- If Honda Accord and BMW 735i has the same market share, a price increase of Toyota Camry has the same impact on them.

Nested Logit Models as Another Special Case

- Group the products into $G+1$ exhaustive and mutually exclusive sets, $g = 0, 1, \ldots, G$.
- Denote the set of products in group g as $\mathcal{J}_g$.
- The outside good, $j = 0$, is the only member of group 0.
- For product $j \in \mathcal{J}_g$, the utility of consumer i is

$$u_{ij} = [x_j\beta - \alpha p_j + \xi_j] + [\zeta_{ig} + (1-\sigma)\varepsilon_{ij}],$$

  - ε_{ij} has an i.i.d extreme value distribution.
  - ζ_{ig} is common to all products in group g and has a distribution function such that $[\zeta_{ig} + (1-\sigma)\varepsilon_{ij}]$ is an extreme value distribution.
  - The parameter $0 \leq \sigma < 1$ measures the within-group substitutability versus the between-group one.

Market Shares under Nested Logit Models

- The market share of product j within group g is

$$\bar{s}_{j/g} = \frac{e^{\frac{\delta_j}{(1-\sigma)}}}{D_g}, \text{ where } \quad D_g \equiv \sum_{j \in \mathcal{J}_g} e^{\frac{\delta_j}{(1-\sigma)}}.$$

- The probability of choosing one of the group g products (the group share) is

$$\bar{s}_g = \frac{D_g^{(1-\sigma)}}{\sum_h D_h^{(1-\sigma)}} = \frac{e^{(1-\sigma)\log D_g}}{\sum_h e^{(1-\sigma)\log D_h}}.$$

- The market share of product j is

$$s_j = \bar{s}_{j/g}\bar{s}_g = \frac{e^{\frac{\delta_j}{(1-\sigma)}}}{D_g^{\sigma}\left[\sum_h D_h^{1-\sigma}\right]}.$$

Linear Representation of the Nested Logit Models

- It can be shown (Berry, 1994 *RAND Journal of Economics*) that

$$\log(s_j) - \log(s_0) = x_j\beta - \alpha p_j + \sigma \log(\bar{s}_{j/g}) + \xi_j.$$

- The coefficients can be estimated from a linear instrumental variables regression.

Random Coefficient Models

- Consumer i's utility of consuming product j can be expressed as

$$u_{ij} = x_j \beta_i - \alpha p_j + \xi_j + \varepsilon_{ij}.$$

(For ease of exposition, assume the price coefficient is constant.)

- Decompose consumer i's taste parameter for characteristic k as

$$\beta_{ik} = \bar{\beta}_k + \sigma_k \zeta_{ik},$$

  - $\bar{\beta}_k$: mean level of the taste parameter for characteristic k
  - σ_k: standard deviation of the taste parameter for k
  - $\zeta_{ik} \sim N(0, 1)$

Alternative Representation of the Utility

- The utility of consuming product j can be written as

$$u_{ij} = \underbrace{[x_j\beta + \xi_j - \alpha p_j]}_{\delta_j} + \Big[\underbrace{\sum_k x_{jk}\sigma_k\zeta_{ik}}_{\mu_{ij}} + \varepsilon_{ij} \Big] = \delta_j + \mu_{ij} + \varepsilon_{ij}.$$

- A consumer selects product j if and only if

$$\delta_j + \mu_{ij} + \varepsilon_{ij} \geq \delta_k + \mu_{ik} + \varepsilon_{ik} \quad \forall k \neq j.$$

- Set of consumer unobservables that lead to the consumption of good j:

$$A_j(\boldsymbol{\delta}) \equiv \{(\boldsymbol{\zeta}_i, \boldsymbol{\varepsilon}_i) | \delta_j + \mu_{ij} + \varepsilon_{ij} \geq \delta_k + \mu_{ik} + \varepsilon_{ik}, \forall k \neq j\}.$$

Market Shares under Random Coefficient Models

- By the law of large numbers, given a distribution $F(\cdot; x, \boldsymbol{\sigma})$ for unobservables $(\boldsymbol{\zeta}, \boldsymbol{\varepsilon})$, the market share of the jth product is

$$s_j(\boldsymbol{\delta}(x, p, \boldsymbol{\xi}), x) = \Pr[(\zeta_i, \varepsilon_i) \in A_j(\boldsymbol{\delta})] = \int_{A_j(\boldsymbol{\delta})} dF(\boldsymbol{\zeta}, \boldsymbol{\varepsilon}; x, \boldsymbol{\sigma}).$$

Market Shares under Random Coefficient Models

- Suppose $\varepsilon_{ij} \sim$ i.i.d. type I extreme value distribution.
- Conditional of the values of ζ, we can compute the market share for product j.

$$f_j(\boldsymbol{\zeta}; x, \boldsymbol{\delta}, \boldsymbol{\sigma}) = \frac{e^{\delta_j + \mu_{ij}(\boldsymbol{\zeta};\boldsymbol{\sigma})}}{\sum_k e^{\delta_k + \mu_{ik}(\boldsymbol{\zeta};\boldsymbol{\sigma})}}.$$

- Let P denote a proper distribution for ζ. The market share of product j predicted by the model is

$$s_j(x, \boldsymbol{\delta}, P, \boldsymbol{\sigma}) = \int f_j(\boldsymbol{\zeta}; x, \boldsymbol{\delta}, \boldsymbol{\sigma}) dP(\boldsymbol{\zeta}) = \int \frac{e^{\delta_j + \mu_{ij}(\boldsymbol{\zeta};\boldsymbol{\sigma})}}{\sum_k e^{\delta_k + \mu_{ik}(\boldsymbol{\zeta};\boldsymbol{\sigma})}} dP(\boldsymbol{\zeta}).$$

- In practice, the above integration is often carried out by simulation.

Inverting the Market Shares into Mean Utility

- In general, we cannot invert the previous equation to get δ.
- BLP propose a recursive method to solve for it numerically.
- Define the operator $T(\mathbf{s}^{obs}, \sigma, P) : \mathbb{R}^J \to \mathbb{R}^J$ such that

$$T(\mathbf{s}^{obs}, \sigma, P)[\boldsymbol{\delta}]_j = \delta_j + \log(s_j^{obs}) - \log(s_j(x, \boldsymbol{\delta}, P, \sigma)),$$

  where s_j^{obs} is the observed market share of product j.
- It can be shown that the operator is a contraction mapping with modulus less than one.
- We can solve for $\boldsymbol{\delta}$ recursively for given observed market shares $\mathbf{s}^{obs}$.

$$\delta_j(\mathbf{s}^{obs}, \sigma) = x_j\beta - \alpha p_j + \xi_j \qquad j = 1, 2, \ldots, J$$

Supply Side: Marginal Costs

- Firms are assumed to be price setters.
- Marginal costs are

$$c_j(q_j, w_j, \omega_j, \gamma),$$

  - w_j: a vector of observed characteristics
  - ω_j: an unobserved characteristic
  - γ: a vector of unknown parameters.
- When the production is constant return to scale, a linear specification of marginal cost is

$$c_j(q_j, w_j, \omega_j, \gamma) = w_j'\gamma + \omega_j.$$

Supply Side: Profit Maximization

- Profits for firm j:

$$\pi_j(p,x,w,\xi,\omega_j,\theta) = p_j M s_j(\delta(x,p,\xi),x,\theta) - TC_j$$

  where the vector θ includes all parameters in the model, and M is population size.

- To maximize profits, the price vector satisfies the usual first-order conditions.

$$M s_j(\delta(x,p,\xi),x,\theta) + p_j M \frac{\partial s_j(\delta(x,p,\xi),x,\theta)}{\partial p_j} - c_j(q_j,w_j,\omega_j,\gamma) M \frac{\partial s_j(\delta(x,p,\xi),x,\theta)}{\partial p_j} = 0.$$

  Therefore,

$$p_j = c_j(q_j,w_j,\omega_j,\gamma) + \left[\frac{\partial s_j(\delta(x,p,\xi),x,\theta)}{\partial p_j}\right]^{-1} s_j(\delta(x,p,\xi),x,\theta).$$

- It's easy to extend the above to consider multi-product firms.

Econometric Assumptions for GMM

- We are going to estimate the parameters by GMM.
- Assume that the supply and demand unobservables are mean independent of both observed product characteristics and cost shifters.

$$E[\xi_j|(x_1, w_1), (x_2, w_2), \ldots, (x_J, w_J)] = 0$$

$$E[\omega_j|(x_1, w_1), (x_2, w_2), \ldots, (x_J, w_J)] = 0$$

- Note: Price or quantity is not included in the conditioning vector.
  - This is because the model implies that price and quantity are determined in part by ξ and ω.
  - The model assumes that product characteristics x_j and cost shifters w_j are exogenous.

Instruments for GMM

- The instruments associated with product j include functions of the characteristics and cost shifters of all other products.
- Products with good substitutes tend to have low markups.
- Because markups will respond differently to own and rival products, the optimal instruments will distinguish between the characteristics of products produced by the same multi-product firm versus the characteristics of products produced by rival firms.
- Let $\mathcal{J}_f$ denote the set of all products produced by firm f.
- BLP's suggestion for the instruments of x_{jk} (the kth characteristic of product j by firm f):

$$x_{jk}, \qquad \sum_{r \neq j, r \in \mathcal{J}_f} x_{rk}, \qquad \sum_{r \notin \mathcal{J}_f} x_{rk}. \qquad \text{(similar IVs for } w_{jk}\text{)}$$

Estimation Algorithm

- Inverting the demand equations gives

$$\delta_j(s^{obs}, \sigma) = x_j\beta - \alpha p_j + \xi_j.$$

- Suppose $H_j(x, w)$ is a matrix of instruments.

$$E\left[H_j(x, w)\begin{pmatrix}\xi_j(s,\theta)\\ \omega_j(s,\theta)\end{pmatrix}\right] = 0$$

- Equivalently, the moment conditions are

$$E\left[H_j(x, w)\begin{pmatrix}\delta_j(s,\theta) - x_j\beta + \alpha p\\ p_j - c_j(q_j, w_j, \gamma) - \frac{1}{\alpha}\left[\frac{s_j}{\partial s_j/\partial\delta_j}\right]\end{pmatrix}\right] = 0.$$

Implement the BLP Estimation Approach

- The goal is to find the GMM estimator for the moment conditions:

$$E\left[H_j(x,w)\begin{pmatrix}\delta_j(s,\theta)-x_j\beta+\alpha p\\ p_j-c_j(q_j,w_j,\gamma)-\frac{1}{\alpha}\left[\frac{s_j}{\partial s_j/\partial\delta_j}\right]\end{pmatrix}\right]=0.$$

- The estimation is carried out as a nested procedure:
  - Outer loop: searching for the minimum of the GMM objective function in the parameter space.
    - "linear" parameters: α, β, γ (easy to obtain by using matrix algebra)
    - "nonlinear" parameters: θ (need numerical methods)
  - Inner loop: For each θ, use contraction mapping to solve for δ_j.

BLP's Original Application

- U.S. automobile market
- 1971–1990
- An observation is a model/year
- Sample size is 2217.
- Estimation Results:
  - Logit benchmark: Table 3 on page 873
  - Full model: Table 4 on page 876
  - Estimated elasticities: Table 6 on page 880
  - Comparison on the substitution to the outside good: Table 7 on page 881

Outline of the Talk

- Background
- Basic Version of the BLP Framework
- **Extensions of the BLP Framework**
- Applications
- Technical Issues and Recent Developments
- Using the BLP Framework in Estimation

Extensions of the BLP Framework

- Adding demographic variables
- Geographic differentiation
- Choice of produce characteristics
- . . .

Adding Demographic Variables

- Nevo (2001, *Econometrica*) combines the market share data with CPS demographic data.
- The coefficients are determined by

$$\begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} = \begin{pmatrix} \bar{\alpha} \\ \bar{\beta} \end{pmatrix} + \Pi D_i + \Sigma \zeta_i,$$

  where D_i is a vector of individual i's demographics, and ζ_i is his unobserved characteristics.
- The estimation procedure is essentially the same.
  - When computing the market share for product j, we also need to take expectations over the distribution of D_i.
  - The distribution of D_i is taken from the empirical distribution observed in the CPS survey.

Another Remark on Nevo (2001)

- Nevo (2001) also considers different instruments (prices in other cities of the region, average city earnings in the supermarket sector)

Geographic Differentiation

- McManus (2007, *RAND Journal of Economics*) studies nonlinear pricing of specialty coffee.
- Consumer's utility is

$$u_{ijlt} = \beta_{xi} q_j^{\gamma_x} + \alpha p_j + \delta D_{jl} + \xi_j + \varepsilon_{ijlt}$$

  - β_{xi}: Consumer i's taste for product line x.
  - γ_x: diminishing utility of consuming product line x.
  - D_{jl}: distance from the consumer's location l to product j.
- Again, we also need to take expectations over the distribution of consumer locations l.

Choice of Produce Characteristics

- Draganska, Mazzeo, and Seim (2009, *Quantitative Marketing and Economics*) consider the choices on the set of products.
- They study the flavors of ice creams.
- In the first stage, the firms decide which flavors to offer.
- In the second stage, firms observe each others' flavor choices. Conditional on their own and their competitors' choice of offerings, firms choose prices.

Outline of the Talk

- Background
- Basic Version of the BLP Framework
- Extensions of the BLP Framework
- **Applications**
- Technical Issues and Recent Developments
- Using the BLP Framework in Estimation

Applications of the BLP Framework

- Price Discrimination
- Two-Sided Markets
- Stockpiling Behavior
- Environmental Policy
- . . .

Price Discrimination

- McManus (2007, *RAND Journal of Economics*) analyzes the pricing scheme of specialty coffee.
- Leslie (2004, *RAND Journal of Economics*) studies price discrimination in Broadway theatre.
- Huang (2008, *Quantitative Marketing and Economics*) considers on-net discounts offered by cellular phone carriers.

Overview of McManus (2007)

- The main goal of the paper is to verify the theoretic prediction about second degree price discrimination: no distortion at the top.
- The author analyzes the market for specialty coffee near the campus of UVa.
- Collect data on prices and sales from all coffee shops in this market for 51 days in Feb./Mar. 2000.
  - There are 9 coffee shops operated by 5 firms.
  - The opening hours differ across stores.
- Cost data are provided by one store manager.
- Population data come from the census and UVa's Registrar, Housing, and Facilities Management offices

Model in McManus (2007)

- Consumer's utility is

$$u_{ijlt} = \beta_{xi} q_j^{\gamma_x} + \alpha p_j + \delta D_{jl} + \xi_j + \varepsilon_{ijlt}$$

  - β_{xi}: Consumer i's taste for product line $x \in \{r, r, s\}$.
    - $\beta_{xi} = \beta_x \exp(\sigma_x v_{xi})$
    - $v_{xi} \sim N(0, 1)$.
  - γ_x: diminishing utility of consuming product line x.
  - D_{jl}: distance from the consumer's location l to product j.
- Conditional on location l, each consumer chooses the highest utility among products available at time t.
- The market share of product j during period t is

$$s_{jt}(\theta) = \int_{A_{jt}} dF_v(v) dF_\varepsilon(\varepsilon) dF_D(D; t)$$

Empirical Results in McManus (2007)

- Figure 1 and Table 1: Industry backgroundd
- Table 5: Parameter estimates
- Figure 2: Distortion in product size

Two-Sided Markets

- Rysman (2004, *Review of Economic Studies*): the market of yellow page directories
- Consumers value directories for information and retailers value directories as a way to advertise to consumers.
- More advertising leads to more consumer usage which in turn leads to more advertising, so consumer behavior and advertiser behavior together create a positive network effect.

Model in Rysman (2004)

- Each publisher $j = 1, 2, \ldots, J$ faces two demand curves: retailer inverse demand for advertising

$$P_j(A_1, U_1, \ldots, A_J, U_J)$$

  and consumer demand for usage

$$U_j(A_1, \ldots, A_J)$$

  where A_j is the amount of advertising at j and U_j is the number of uses per consumer covered.
- Because of network effects, expect $\partial P_j / \partial U_j > 0$ and $\partial U_j / \partial A_j > 0$. Law of demand suggests that $\partial P_j / \partial U_j < 0$.
- Each publisher simultaneously chooses its quantity of advertising A_j to sell in its directory.

Econometric Specification

- The advertiser's inverse demand is specifies as

$$\ln(P_j) = \gamma \ln A_j + \alpha_1 \ln U_j + X_j^P \beta^P + \nu_j .$$

- Let the utility to consumer i from directory j be

$$u_{ij} = \alpha_2 \ln A_j + X_j^U \beta^U + \xi_j + \underbrace{\sigma\zeta_i + (1-\sigma)\varepsilon_{ij}}_{\text{nested logit error terms}}$$

where ζ_i and ε_{ij} are assumed to have the distribution for the nested logit model.

- We will have

$$\ln s_j - \ln s_0 = \alpha_2 \ln A_j + X_j^U \beta^U + \sigma \ln(s_j|\text{Yellow Page}) + \xi_j$$

Econometric Specification

- The profit maximization problem is

$$\max_{A_j} \sum_{k \in K(j)} P_k(A_k, U_k(A_1, \ldots, A_J))A_K - MC_j A_j$$

- Assume the marginal cost being

$$MC_j = \exp(X_j^C \beta^C + \omega_j)$$

- We can compute the first order condition for each publisher.
- Estimate the model by GMM.

Estimation Results

- Table 5: Parameter estimates
- Policy simulation on the effects of entry:
  - Despite network effects, welfare increases in the number of directories.
  - Figure 1: The current distribution on the number of yellow pages.
  - Table 7 and 8: Simulation of entries
  - Figure 4: Larger network effects

Application: Stockpiling

- Hendel and Nevo (2006, *Econometrica*) estimate a dynamic model of consumer choice to account for intertemporal substitution.
- They study the effect of sales (temporary price reductions) on consumer behavior.
- Policy analysis based on static elasticity estimates will underestimate price-cost margins and underpredict the effects of mergers.
- Data: two years of scanner data on the purchase of laundry detergents.

Consumer Model of Hendel and Nevo (2006)

- Consumer h obtains per period utility:

$$u(c_{ht}, \nu_{ht}; \theta_t) + \alpha_h m_{ht}$$

  - c_{ht}: consumption of detergents
  - ν_{ht}: idiosyncratic demand shocks
  - m_{ht}: consumption of the outside good

- x_{ht}: the purchased amount of detergents (four types: 32, 64, 96, or 128 oz.)
- $d_{hjxt} \in \{0, 1\}$: a purchase of brand j and size x (Assume $\sum_{j,x} d_{hjxt} = 1$.)

Consumer's Dynamic Choice Problem

$$V(s_1) = \max_{\{c_h(s_t), d_{hjx}(s_t)\}} \sum_{t=1}^{\infty} \delta^{t-1} E\Big[u(c_{ht}, \nu_{ht}; \theta_g) - C_h(i_{h,t+1}; \theta_h)$$

$$+ \sum_j d_{hjxt}(\alpha_h p_{jxt} + \xi_{hjt} + \beta_h a_{jxt} + \varepsilon_{hjxt}) | s_1 \Big]$$

$$\text{s.t. } 0 \le i_{ht}, 0 \le c_{ht}, 0 \le x_{ht}, \sum_{jx} d_{hjxt} = 1, i_{h,t+1} = i_{ht} + x_{ht} - c_{ht}.$$

- s_t: state at time t
- $C_h(i; \theta_h)$: cost of storage with inventory i
- ξ_{hjt}: idiosyncratic taste for brand j
- $\beta_h a_{jxt}$: effect of advertising
- ε_{hjxt}: random shock $\sim$ type-I extreme value

Three-Step Estimation Procedure

- Step 1: Estimation of the "static" parameters
  - The optimal consumption, conditional on the quantity purchased, is not brand-specific.

$$\Pr(d_{jt} = 1 | x_t, i_t, p_t, \nu_t) = \Pr(d_{jt} = 1 | x_t, p_t).$$

- Step 2: Compute the inclusive values (a consumer's expected surplus of purchasing size x).

$$\begin{aligned} \omega_{xt} &\equiv E\left[\max_k \{\alpha p_{kxt} + \xi_{kxt} + \beta a_{kxt} + \varepsilon_{kxt}\}\right] \\ &= \log\left[\sum_k \exp(\alpha p_{kxt} + \xi_{kxt} + \beta a_{kxt})\right] \end{aligned}$$

- Step 3: Estimate Simplified Dynamic Problem on c and x.

Empirical Results

- Table 4: Brand choice conditional on size
- Table 6: Estimates of dynamic parameters
- Figure 1: Hazard rate of purchases
- Table 8: Comparison between the short-run and long-run elasticities

Application: Environmental Policy

- Beresteanu and Li (2011, *International Economic Review*) estimate the demand for hybrid vehicles to evaluate the effects of environmental policies.
- Government policies:
    - federal tax deductions
    - federal tax credits
    - local incentives (income tax, excise tax, sales tax, HOV)

Estimation Approach in Beresteanu and Li (2011)

- The model is similar to BLP (1995).
- Sales data at MSA level for 22 MSAs in 1999–2006.

Results in Beresteanu and Li (2011)

- Table 5: parameter estimates
- Table 8 and 9: effects of gasoline price rise
- Table 10 and 11: effects of federal tax incentives
- Table 12: comparison of policies

Outline of the Talk

- Background
- Basic Version of the BLP Framework
- Extensions of the BLP Framework
- Applications
- **Technical Issues and Recent Developments**
- Using the BLP Framework in Estimation

Technical Issues and Recent Developments

- Dubé, Fox, and Su (2012, *Econometrica*):
- Knittel and Metaxoglou (2014, *Review of Economics and Statistics*):

Numerical Performance of the BLP Estimator

- The algorithm proposed in BLP is a nested fixed point (NFP) approach.
  - Outer loop: optimization over structural parameters
  - Inner loop: a contraction mapping
- Dubé, Fox, and Su (2012): It is crucial to use a stringent stopping criterion in both the outer and the inner loops of the NFP approach.

Biases Caused by Loose Tolerances

- See Table 1 and Table 2 of Dubé, Fox, and Su (2012).
- Knittel and Metaxoglou (2014) also show that the numerical details (optimization algorithm, tolerance levels, starting points) in implementing the BLP framework can serious affect the estimation results.
- It is important to report the numerical details.

The MPEC Approach

- Dubé, Fox, and Su (2012) propose an alternative numerical approach: a mathematical program with equilibrium constraints (MPEC).
- The BLP estimator can be expressed as

$$\min_{\theta,\xi} g(\xi)' W g(\xi) \qquad \text{s.t. } s(\xi;\theta) = S,$$

where W is the GMM weighting matrix, and the moment condition is

$$g(\xi) = \frac{1}{T}\sum_{t=1}^{T}\sum_{j=1}^{J}\xi_{jt} g(z_{jt}, x_{jt})$$

The MPEC Approach: Alternative Representation

- An alternative formulation of the MPEC approach:

$$\min_{\theta,\xi,\eta} \eta' W \eta \qquad \text{s.t. } g(\xi) = \eta$$
$$s(\xi; \theta) = S,$$

- Advantages of the MPEC approach:
  - no numerical errors from the inner loop
  - faster speed

Outline of the Talk

- Background
- Basic Version of the BLP Framework
- Extensions of the BLP Framework
- Applications
- Technical Issues and Recent Developments
- **Using the BLP Framework in Estimation**

Logit and Nested-Logit Models

- Logit and nested logit models are linear in unknown parameters.
- They can be easily executed by standard softwares, such as Stata.

Random Coefficient Models

- In practice, we want to do GMM estimation for a set of nonlinear moment conditions.
- I usually use Matlab.
- A simple example: the utility of consumer i on choosing product j is

$$u_{ij} = \beta_0 + \beta_{1i} x_{1j} + \beta_2 x_{2j} + \beta_3 x_{3j} + \varepsilon_{ij}$$

  with
  - $\beta_{1i} \sim N(\beta_1, \sigma^2)$
  - $\varepsilon_{ij} \sim$ type-1 extreme value