Graduate IO: Demand Estimation

October 22, 2017

Motivation

- much of the information required to analyze the incentives firms facing in a market

Motivation

- much of the information required to analyze the incentives firms facing in a market
- these incentives play a major role in determining profits to be earned
  - different pricing decisions
  - alternative investments: new products, advertising (how to regulate it?)

Motivation

- much of the information required to analyze the incentives firms facing in a market
- these incentives play a major role in determining profits to be earned
    - different pricing decisions
    - alternative investments: new products, advertising (how to regulate it?)
- typical empirical questions: response of prices and product development due to policy or environmental changes

Motivation (Cont'd.)

- many other components determining the profitability of pricing and product development, but typically hard to analyze
  - static response to a change in the environment: cost function, equilibrium assumption
    - cost data are typically proprietary
    - analyzing the nature of competition: less progress using equilibrium assumptions than demand systems

Welfare Analysis

- demand analysis framework: aggregated from individual specific utility functions and choices to market level
  - pricing or product placement decisions could be discussed
  - analysis of distributional impacts on utility resulted from policy or environmental changes

Welfare Analysis

- demand analysis framework: aggregated from individual specific utility functions and choices to market level
  - pricing or product placement decisions could be discussed
  - analysis of distributional impacts on utility resulted from policy or environmental changes
- consumer surplus gains to product introductions, e.g., Minivan (Petrin 2002 JPE), PC (Hendel 1999 REStud), cloud computing
  - determine society's gain from private R&D activity (which is often subsidized)
  - evaluate the effectiveness of the proprietary rights established to foster inventive and creative activity (patents, copyright, trademarks)

Regulatory Policies

- pricing policies: regulated utilities, e.g., water pricing (Timmins 2002 ECMA)

Regulatory Policies

- pricing policies: regulated utilities, e.g., water pricing (Timmins 2002 ECMA)
- decisions to either foster or limit the marketing of goods and services in regulated markets: advertising restrictions, regulatory delay, auctioning off the spectrum
  - regulations on content of advertising: FTC
  - spectrum and telecommunication rules: FCC

Regulatory Policies

- pricing policies: regulated utilities, e.g., water pricing (Timmins 2002 ECMA)
- decisions to either foster or limit the marketing of goods and services in regulated markets: advertising restrictions, regulatory delay, auctioning off the spectrum
  - regulations on content of advertising: FTC
  - spectrum and telecommunication rules: FCC
- note: regulatory decisions are motivated by non-market factors, e.g., insure access to particular services to all members of community, political requirements
  - cannot only focus on mean/median: need to evaluate the full distribution of implications/impacts
  - seems quite useful in China because of heavy regulations in many industries

Single-Product Demand Function: Simultaneity Problem

- assume an iso-elastic (log-log) demand

$$\log(q_{jt}) = \alpha_j \log(p_{jt}) + X_{jt}\beta + \xi_{jt}$$

Single-Product Demand Function: Simultaneity Problem

- assume an iso-elastic (log-log) demand

$$\log(q_{jt}) = \alpha_j \log(p_{jt}) + X_{jt}\beta + \xi_{jt}$$

- price endogeneity: suppose ξ is correlated with $\log(p)$, e.g., $E(\xi_{jt}\log(p_{jt})) \neq 0$

$$\xi_{jt} = \lambda \log(p_{jt}) + \varepsilon_{jt},\ E[\varepsilon_{jt}\log(p_{jt})] = 0$$

  so that we can re-write the demand function as

$$\log(q_{jt}) = (\alpha_j + \lambda)\log(p_{jt}) + X_{jt}\beta + \varepsilon_{jt}$$

- the estimated α will be biased upwards (downwards) if λ is positive (negative)
  - typically cannot get experimental data in IO
  - unlike econometrics, we need to think about what goes into the "unobservables" (instead of just an "error term")
  - it is important to recognize that the firm will react to the unobservable: we do not have data does NOT mean the firm does not either

Demand System

- goal: estimate markups, this requires estimates of demand and supply

Demand System

- goal: estimate markups, this requires estimates of demand and supply
- suppose there are J products, the demand system is given by

$$q = D(p, x)$$

where

  - q is a J-dimensional vector of quantities demanded
  - p is J dimensional vector of prices
  - x is a vector of demand shifters

Demand System

- goal: estimate markups, this requires estimates of demand and supply
- suppose there are J products, the demand system is given by

  $$q = D(p, x)$$

  where

  - q is a J-dimensional vector of quantities demanded
  - p is J dimensional vector of prices
  - x is a vector of demand shifters
- need to specify a functional form for the demand system that is consistent with choice theory and flexible enough to fit the data
  - examples: translog, Almost Ideal Demand System (AIDS), linear expenditure system, etc

Dimensionality Problem

- basic problem in IO (in contrast to labor, transportation, etc.): number of alternatives J is typically quite large
  - number of beer brands are at least 50
  - number of models of cars is over 100
  - number of houses/apartments is over 10,000

Dimensionality Problem

- basic problem in IO (in contrast to labor, transportation, etc.): number of alternatives J is typically quite large
  - number of beer brands are at least 50
  - number of models of cars is over 100
  - number of houses/apartments is over 10,000
- the number of parameters to be estimated (i.e., own/cross price elasticities) is on the order of J^2 - way too many parameters

Solutions

- focus on an aggregate (e.g., Porter aggregates all eastbound shipments rather than differentiating across destination cities) or on a small subset of the products (Baker and Bresnahan study a particular segment of the beer industry)
  - basic question: do you need to estimate the full demand system to answer the question in which you are interested?

Solutions

- focus on an aggregate (e.g., Porter aggregates all eastbound shipments rather than differentiating across destination cities) or on a small subset of the products (Baker and Bresnahan study a particular segment of the beer industry)
  - basic question: do you need to estimate the full demand system to answer the question in which you are interested?
- product space model: impose structure on preferences such as symmetry or separability to restrict the substitution patterns across products

Solutions

- focus on an aggregate (e.g., Porter aggregates all eastbound shipments rather than differentiating across destination cities) or on a small subset of the products (Baker and Bresnahan study a particular segment of the beer industry)
  - basic question: do you need to estimate the full demand system to answer the question in which you are interested?

- product space model: impose structure on preferences such as symmetry or separability to restrict the substitution patterns across products

- characteristic approach: define products as bundles of a limited number of characteristics and define preferences on characteristics rather than products

Dixit-Stiglitz or CES Model

- the CES utility function

$$U(q_1, ..., q_J) = \left(\sum_{j=1}^{J} q_j^{\sigma} \right)^{1/\sigma}$$

where σ measures rate of substitution across products

Dixit-Stiglitz or CES Model

- the CES utility function

$$U(q_1, ..., q_J) = \left(\sum_{j=1}^{J} q_j^{\sigma} \right)^{1/\sigma}$$

where σ measures rate of substitution across products

- maximizing U subject to a linear budget constraint yields demand

$$q_j = \frac{p_j^{-1/(1-\sigma)} I}{\sum_{k=1}^{J} p_k^{-\sigma/(1-\sigma)}}, \; j = 1, ..., J$$

where I is consumer income

Dixit-Stiglitz or CES Model

- the CES utility function

$$U(q_1, \ldots, q_J) = \left(\sum_{j=1}^{J} q_j^{\sigma} \right)^{1/\sigma}$$

  where σ measures rate of substitution across products

- maximizing U subject to a linear budget constraint yields demand

$$q_j = \frac{p_j^{-1/(1-\sigma)} I}{\sum_{k=1}^{J} p_k^{-\sigma/(1-\sigma)}}, \; j = 1, \ldots, J$$

  where I is consumer income

- cost of using CES is that it imposes strong and implausible restrictions on own and cross price elasticities

$$\frac{\partial q_i}{\partial p_j} \frac{p_j}{q_i} = \frac{\partial q_k}{\partial p_j} \frac{p_j}{q_k}, \text{ for all } i, j, k$$

  - rules out differential substitution, like simple logit (but even

More General Approach: AIDS in Deaton and Mullbauer (1980)

- divide products into smaller groups and allow for flexible functional form within each group

More General Approach: AIDS in Deaton and Mullbauer (1980)

- divide products into smaller groups and allow for flexible functional form within each group
- two ideas: separability and multi-stage budgeting

More General Approach: AIDS in Deaton and Mullbauer (1980)

- divide products into smaller groups and allow for flexible functional form within each group
- two ideas: separability and multi-stage budgeting
- (weakly) separable preferences: q can be partitioned into $\left(q^1, \ldots, q^N\right)$

$$U = f\left(v_1\left(q^1\right), \ldots, v_N\left(q^N\right)\right)$$

where $v_K\left(q^K\right)$ is a sub-utility function (i.e., represents a preference ordering over q^K) and f is an increasing function in all of its arguments

Two-Step Optimization

- implication: maximizing U subject to linear budget constraint is equivalent to following two step optimization program
  - step 1: fix allocation of income across the commodity groups $\left(I^1, \ldots, I^N\right)$ and solve N optimization problems of the form

    $$\max v_K\left(q^K\right) \ s.t. \ p^K q^K = I^K, \ K = 1, .., N$$

    the solution to each of these subproblems is a set of subgroup demands of the form

    $$q^K = g^K\left(p^K, I^K\right)$$

  - step 2: substitute subgroup demands into subgroup utility functions to obtain the indirect utility functions $\psi^K\left(p^K, I^K\right)$, the choose allocation of income to solve

    $$\max_{\left(I^1,\ldots,I^N\right)} f\left(\psi^1\left(p^1, I^1\right), ..., \psi^N\left(p^N, I^N\right)\right) \ s.t. \ \sum_{K=1}^{N} I^K = x$$

Separable Preferences

- weak separability is a necessary and sufficient condition for the first step
  - preferences are typically assumed to be additively (strong) separable over time
  - preferences over goods are typically assumed to be separable from leisure

Separable Preferences

- weak separability is a necessary and sufficient condition for the first step
  - preferences are typically assumed to be additively (strong) separable over time
  - preferences over goods are typically assumed to be separable from leisure
- separability implies a reduction in the number of parameters
  - suppose each group consists of m products, number of parameters is $Nm^2/2$ rather than $(Nm)^2/2$
  - it can be tested because it imposes restrictions on the substitution matrix

Multi-stage Budgeting

- in the second step, we often would like to treat the optimization problem as one in which the consumer chooses quantities of the composite commodity groups to maximize utility

Multi-stage Budgeting

- in the second step, we often would like to treat the optimization problem as one in which the consumer chooses quantities of the composite commodity groups to maximize utility

- for example, consumer demands for food, clothing, shelter, and entertainment are often expressed as functions of price indices for these commodities and income

Application: Hausman Papers on Demand in Beer and Cereal Market

- Hausman, Leonard and Zona (1994) and Hausman (1996)

Application: Hausman Papers on Demand in Beer and Cereal Market

- Hausman, Leonard and Zona (1994) and Hausman (1996)
- data: brand j's prices and shares, by city c and quarter t

Application: Hausman Papers on Demand in Beer and Cereal Market

- Hausman, Leonard and Zona (1994) and Hausman (1996)
- data: brand j's prices and shares, by city c and quarter t
- multi-level demand system with three levels
  - top level: overall demand for the product, e.g., beer or ready-to-eat cereal
  - middle level: demand for different market segments, e.g., in beer, lager, pilsner and ale; in cereals, family, kids and adult cereals
  - bottom level: a system of demands for different brands in each segment

Bottom Level

- demand for brand j within segment g in city c in quarter t is

$$s_{jct} = \alpha_{jc} + \beta_j \log\left(\frac{I_{gct}}{P_{gct}}\right) + \sum_{k=1}^{J_g} \gamma_{jk} \log\left(p_{kct}\right) + \varepsilon_{jct}$$

where s_{jct} is segment expenditure share of brand j

Bottom Level

- demand for brand j within segment g in city c in quarter t is

$$s_{jct} = \alpha_{jc} + \beta_j \log\left(\frac{I_{gct}}{P_{gct}}\right) + \sum_{k=1}^{J_g} \gamma_{jk} \log\left(p_{kct}\right) + \varepsilon_{jct}$$

  where s_{jct} is segment expenditure share of brand j

- here P_{gct} is the price index for segment g in city c in quarter t and J_g is the number of brands in segment g
  - Stone logarithmic price index

$$P_{gct} = \sum_{j \in g} s_{jct} \log\left(p_{jct}\right)$$

  - Deaton and Mullbauer exact price index

$$P_{gct} = \alpha_0 + \sum_{j \in g} \alpha_j p_{jct} + \sum_{j \in g} \sum_{k \in g} \gamma_{jk} \log\left(p_{kct}\right) \log\left(p_{jct}\right)$$

Bottom Level

- demand for brand j within segment g in city c in quarter t is

$$s_{jct} = \alpha_{jc} + \beta_j \log\left(\frac{I_{gct}}{P_{gct}}\right) + \sum_{k=1}^{J_g} \gamma_{jk} \log\left(p_{kct}\right) + \varepsilon_{jct}$$

  where s_{jct} is segment expenditure share of brand j

- here P_{gct} is the price index for segment g in city c in quarter t and J_g is the number of brands in segment g
  - Stone logarithmic price index

$$P_{gct} = \sum_{j \in g} s_{jct} \log\left(p_{jct}\right)$$

  - Deaton and Mullbauer exact price index

$$P_{gct} = \alpha_0 + \sum_{j \in g} \alpha_j p_{jct} + \sum_{j \in g} \sum_{k \in g} \gamma_{jk} \log\left(p_{kct}\right) \log\left(p_{jct}\right)$$

- with stone price index, brands demand can be estimated using linear methods; the Deaton and Mullbauer price index requires non-linear methods

Middle and Top Level Demands

- middle level

$$\log\left(q_{gct}\right) = \beta_g \log\left(I_{ct}\right) + \sum_{g=1}^{G} \delta_g \log\left(\pi_{gct}\right) + \alpha_{gc} + \varepsilon_{gct}$$

where q_{gct} is the composite quantity of the g segment in city c in quarter t, π_{gct} are the segment price indices computed as above

Middle and Top Level Demands

- middle level

$$\log(q_{gct}) = \beta_g \log(I_{ct}) + \sum_{g=1}^{G} \delta_g \log(\pi_{gct}) + \alpha_{gc} + \varepsilon_{gct}$$

where q_{gct} is the composite quantity of the g segment in city c in quarter t, π_{gct} are the segment price indices computed as above

- top level

$$\log(q_{ct}) = \beta_0 + \beta_1 \log(I_{ct}) + \beta_2 \log(\pi_{ct}) + \theta Z_{ct} + \varepsilon_{ct}$$

where q_{ct} is the quantity of beer (or cereal) in city c in quarter t, I_{ct} is the expenditure on beer in city$c$ in quarter t, π_{ct} is the price index for beer, and Z_{ct} are demand shifters

Another Application: Hausman (1997)

- issue: regulatory delay in the introduction of telecommunications innovations

Another Application: Hausman (1997)

- issue: regulatory delay in the introduction of telecommunications innovations
- in 1947, Bell Labs (the research arm of) AT&T worked out the technology that makes mobile phones practical

Another Application: Hausman (1997)

- issue: regulatory delay in the introduction of telecommunications innovations
- in 1947, Bell Labs (the research arm of) AT&T worked out the technology that makes mobile phones practical
- AT&T proposed that FCC allocate a large number of radio-spectrum frequencies for mobile phone service

Another Application: Hausman (1997)

- issue: regulatory delay in the introduction of telecommunications innovations
- in 1947, Bell Labs (the research arm of) AT&T worked out the technology that makes mobile phones practical
- AT&T proposed that FCC allocate a large number of radio-spectrum frequencies for mobile phone service
- FCC decided to limit frequency availability, AT&T slowed its investment in mobile technology

Regulatory Delay

- in 1968, FCC reconsidered, if the technology became available they would increase the spectrum allocation
  - by 1977 AT&T Bell Labs and then Motorola had a prototype cellular system

Regulatory Delay

- in 1968, FCC reconsidered, if the technology became available they would increase the spectrum allocation
  - by 1977 AT&T Bell Labs and then Motorola had a prototype cellular system
- FCC authorize the use of cellular service in 1982 and the first system were setup in 1983 and 1984
  - meanwhile, NTT in Japan had a citywide cellular network launched in 1979, Denmark, Finland, Sweden, Norway had systems in 1981

Regulatory Delay

- in 1968, FCC reconsidered, if the technology became available they would increase the spectrum allocation
  - by 1977 AT&T Bell Labs and then Motorola had a prototype cellular system
- FCC authorize the use of cellular service in 1982 and the first system were setup in 1983 and 1984
  - meanwhile, NTT in Japan had a citywide cellular network launched in 1979, Denmark, Finland, Sweden, Norway had systems in 1981
- US corporations has lost much of the international market by the time they start

Why the FCC Delayed So Long?

- maybe they worried about the competitive effects of granting AT&T a mobile phone license at the same time that it had monopoly rights over hard line technology

Why the FCC Delayed So Long?

- maybe they worried about the competitive effects of granting AT&T a mobile phone license at the same time that it had monopoly rights over hard line technology
- the problem of course was that they did not take into account the costs of such delay and they were very large

Why the FCC Delayed So Long?

- maybe they worried about the competitive effects of granting AT&T a mobile phone license at the same time that it had monopoly rights over hard line technology
- the problem of course was that they did not take into account the costs of such delay and they were very large
- Hausman estimates the consumer surplus gain of introducing cell phone, which is about 30 billion dollars a year, and it does not count the losses which resulted from not marketing earlier in foreign countries

Discrete Choice Models

- products are bundles of characteristics: $j = 0, .., J$ where 0 is the outside good

Discrete Choice Models

- products are bundles of characteristics: $j = 0, .., J$ where 0 is the outside good
- consumer preferences are defined on space of characteristics

Discrete Choice Models

- products are bundles of characteristics: $j = 0, .., J$ where 0 is the outside good
- consumer preferences are defined on space of characteristics
- each consumer $i = 1, .., N$ chooses at most one unit of one of the inside goods, the choice maximizes utility

Discrete Choice Models

- products are bundles of characteristics: $j = 0, .., J$ where 0 is the outside good
- consumer preferences are defined on space of characteristics
- each consumer $i = 1, .., N$ chooses at most one unit of one of the inside goods, the choice maximizes utility
- consumers are heterogeneous: they have different preferences for different characteristics, the distribution of heterogeneity is parameterized

Discrete Choice Models (Cont.)

- in this model, we are estimating the joint distribution of preferences over characteristics
  - the number of parameters is primarily determined by the dimensionality of the characteristics and independent of the number of products

Discrete Choice Models (Cont.)

- in this model, we are estimating the joint distribution of preferences over characteristics
  - the number of parameters is primarily determined by the dimensionality of the characteristics and independent of the number of products
- the parameter estimates measure the consumers' marginal valuation of characteristics: if a new good is introduced, we can value that good since it is simply a bundle of characteristics
  - we can also predict outcomes demand for the old and new goods in the expanded choice set

Discrete Choice Models (Cont.)

- in this model, we are estimating the joint distribution of preferences over characteristics
  - the number of parameters is primarily determined by the dimensionality of the characteristics and independent of the number of products
- the parameter estimates measure the consumers' marginal valuation of characteristics: if a new good is introduced, we can value that good since it is simply a bundle of characteristics
  - we can also predict outcomes demand for the old and new goods in the expanded choice set
- caveat: the new good cannot be too "new" – i.e., possess a new characteristic

Basic Model

- utility of consumer i for product j is given by

$$u_{ij} = U\left(x_j, p_j, \xi_j, \upsilon_i; \theta\right)$$

  where

  - x_j: a vector of observed characteristics of product j
  - ξ_j: unobserved characteristic of product j
  - p_j: price of product j
  - $\upsilon_i \sim F$: unobserved preference characteristics of consumer i
  - θ: vector of utility parameters to be estimated

Choice Probability

- define the subset of "consumers" (preferences) that lead to choice j as

$$A_j(\theta) = \{\upsilon \,|\, u_{ij} > u_{ik}, \forall k\}$$

Choice Probability

- define the subset of "consumers" (preferences) that lead to choice j as
$$A_j(\theta) = \{\upsilon \,|\, u_{ij} > u_{ik}, \forall k\}$$

- then the probability that consumer i chooses product j is
$$\sigma_j(x, p, \xi; \theta) = \int_{\upsilon \in A_j(\theta)} f(\upsilon)\, d\upsilon$$

where $x = (x_1, ..., x_J)$, $p = (p_1, ..., p_J)$ and f is the density associated with F

Choice Probability

- define the subset of "consumers" (preferences) that lead to choice j as

$$A_j(\theta) = \{\upsilon \,|\, u_{ij} > u_{ik}, \forall k\}$$

- then the probability that consumer i chooses product j is

$$\sigma_j(x, p, \xi; \theta) = \int_{\upsilon \in A_j(\theta)} f(\upsilon)\, d\upsilon$$

where $x = (x_1, ..., x_J)$, $p = (p_1, ..., p_J)$ and f is the density associated with F

- under the assumption that "market size" M is very large and υ_i's are i.i.d. across consumers, the Law of Large Numbers implies that market demand converges to $M\sigma_j(x, p, \xi; \theta)$

Remarks

- if there is no outside good, then the market is covered and aggregate demand is M, the number of consumers in the market $\rightarrow$ inelastic market demand, no market expansion effects

Remarks

- if there is no outside good, then the market is covered and aggregate demand is M, the number of consumers in the market $\rightarrow$ inelastic market demand, no market expansion effects
- normalizations: choices of individual consumers are invariant to affine transformation of utilities
  - invariance to additive shifts implies normalizing mean utility of outside good to zero $\rightarrow$ deduct u_{i0} from each u_{ij} for $j = 1, .., J$
  - invariance to scale leads to normalizing one of the other parameters (typically variance of F) to one

Generation I Models

- data: $\{s_j, p_j, x_j\}$ where s_j is the observed market share of product j

Generation I Models

- data: $\{s_j, p_j, x_j\}$ where s_j is the observed market share of product j

- basic idea is to estimate θ (which includes the parameters of F) by minimizing the distance between the predicted choice probabilities and observed market shares
  - the choice model determines $(\sigma_0(\theta), \sigma_1(\theta), ..., \sigma_J(\theta))$
  - each consumer is an independent draw from F, then the distribution of product choices is given by a multinomial distribution

Estimation

- let q_j denote the number of consumers who choose product j, the likelihood function for the data is

$$L = \Pi_{j=0}^{J} \left[\sigma_j\left(\theta\right)\right]^{q_j}$$

Estimation

- let q_j denote the number of consumers who choose product j, the likelihood function for the data is

$$L = \Pi_{j=0}^{J} \left[\sigma_j(\theta)\right]^{q_j}$$

- taking logs, choose θ to

$$\max_{\theta} M \sum_{j=0}^{J} s_j \log\left[\sigma_j(\theta)\right] \Leftrightarrow \min_{\theta} \sum_{j=0}^{J} \frac{\left[s_j - \sigma_j(\theta)\right]^2}{\sigma_j(\theta)}$$

  - last statement follows from taking a Taylor series approximation of $\sigma_j(\theta)$ around the data point s_j
  - the latter statistic is called a minimum χ^2, if we use observed shares in denominator, then it is modified χ^2

Example: Logit Model

- utility function

$$u_{ij} = x_j \beta - p_j + \epsilon_{ij}$$

  where ϵ_{ij} is distributed i.i.d. with mean zero across products and consumers and its distribution is Type I extreme value, i.e., $F(\epsilon) = \exp[-\exp(-\epsilon)]$

Example: Logit Model

- utility function

$$u_{ij} = x_j\beta - p_j + \epsilon_{ij}$$

where ϵ_{ij} is distributed i.i.d. with mean zero across products and consumers and its distribution is Type I extreme value, i.e., $F(\epsilon) = \exp\left[-\exp\left(-\epsilon\right)\right]$

- predicted choice probabilities (market shares)

$$\sigma_j(\theta) = \frac{\exp\left(x_j\beta - p_j\right)}{\sum_{k=0}^{J} \exp\left(x_k\beta - p_k\right)},\ j = 0, 1, ..., J$$

  - remark: McFadden (1974) shows that the multinomial logit model is derived from utility maximization if and only if $\{\epsilon_{ij}\}$ are independent across products and distributed Type I extreme value

Example: Logit Model (Cont.)

- normalize the mean utility of the outside good to zero implies that

$$\sigma_0(\theta) = \frac{1}{\sum_{k=0}^{J} \exp(x_k \beta - p_k)}$$

Example: Logit Model (Cont.)

- normalize the mean utility of the outside good to zero implies that
$$\sigma_0(\theta) = \frac{1}{\sum_{k=0}^{J} \exp(x_k \beta - p_k)}$$

- match predicted market shares to observed ones: the model is correctly specified

$$s_j = \sigma_j(\theta), \ j = 0, 1, ..., J$$

Example: Logit Model (Cont.)

- normalize the mean utility of the outside good to zero implies that
$$\sigma_0(\theta) = \frac{1}{\sum_{k=0}^{J} \exp(x_k \beta - p_k)}$$
- match predicted market shares to observed ones: the model is correctly specified
$$s_j = \sigma_j(\theta),\ j = 0, 1, ..., J$$
- therefore
$$\log(s_j) - \log(s_0) = x_j \beta - p_j$$
  - no additional taste parameter so $\theta = \beta$

Vertical Model: Bresnahan (1981)

- utility function

$$u_{ij} = \upsilon_i \varphi_j - p_j,\ \upsilon_i > 0$$

where φ_j measures the quality of good j and is assumed to be strictly increasing in j

  - no unobserved product characteristics: $\varphi_j = x_j \beta$

Vertical Model: Bresnahan (1981)

- utility function
$$u_{ij} = \upsilon_i \varphi_j - p_j, \ \upsilon_i > 0$$
where φ_j measures the quality of good j and is assumed to be strictly increasing in j
  - no unobserved product characteristics: $\varphi_j = x_j \beta$

- product choice probabilities (product demands)
  - necessary condition of positive demand
$$\upsilon \varphi_j - p_j > \upsilon \varphi_{j+1} - p_{j+1}$$
$$\upsilon \varphi_j - p_j > \upsilon \varphi_{j-1} - p_{j-1}$$
which implies
$$\frac{p_j - p_{j-1}}{\varphi_j - \varphi_{j-1}} < \upsilon < \frac{p_{j+1} - p_j}{\varphi_{j+1} - \varphi_j}$$

Market Shares

- define

$$\Delta_j = \frac{p_j - p_{j-1}}{(x_j - x_{j-1})\,\beta},\ J > j > 0$$

and $\Delta_0 = -\infty$, $\Delta_J = \infty$

Market Shares

- define
$$\Delta_j = \frac{p_j - p_{j-1}}{(x_j - x_{j-1})\,\beta},\ J > j > 0$$
and $\Delta_0 = -\infty$, $\Delta_J = \infty$

- necessary and sufficient condition for demands for all J goods to be positive is that Δ_j is strictly increasing in j

Market Shares

- define
$$\Delta_j = \frac{p_j - p_{j-1}}{(x_j - x_{j-1})\,\beta},\ J > j > 0$$
and $\Delta_0 = -\infty$, $\Delta_J = \infty$

- necessary and sufficient condition for demands for all J goods to be positive is that Δ_j is strictly increasing in j

- market share of product j is
$$\sigma_j(\theta) = F(\Delta_{j+1}) - F(\Delta_j)$$
where F is the distribution of υ

Remarks

- normalizations: (i) $\varphi_0 = 0$, (ii) $p_0 = 0$

Remarks

- normalizations: (i) $\varphi_0 = 0$, (ii) $p_0 = 0$
- here $\theta = (\beta, \lambda)$ where λ is the parameter of F: choose θ to minimize the difference between the observed and predicted market shares

Remarks

- normalizations: (i) $\varphi_0 = 0$, (ii) $p_0 = 0$
- here $\theta = (\beta, \lambda)$ where λ is the parameter of F: choose θ to minimize the difference between the observed and predicted market shares
- differences between actual market shares and the choice probabilities can only be due to sampling error
  - as $M \to \infty$, $s_j \to \infty$, model should fit exactly
  - lack of prediction error means model is certain to be rejected by the data: no value for β such that actual shares $=$ predicted shares

Generation II: Berry (1994)

- theory of the error is unobserved product characteristics: ξ_j

Generation II: Berry (1994)

- theory of the error is unobserved product characteristics: ξ_j
- mean utility for product j is

$$\delta_j = x_j\beta - \alpha p_j + \xi_j$$

Generation II: Berry (1994)

- theory of the error is unobserved product characteristics: ξ_j
- mean utility for product j is

$$\delta_j = x_j\beta - \alpha p_j + \xi_j$$

- vertical model is a special case: $\varphi_j = x_j\beta + \xi_j$, $\alpha = 1$ and the mean of υ is normalized to 1

Endogeneity

- main problem: firms know ξ when they set prices so prices are correlated with the error which in turn is buried in a highly non-linear set of equations

Endogeneity

- main problem: firms know ξ when they set prices so prices are correlated with the error which in turn is buried in a highly non-linear set of equations

- solution: assume M is large enough that

$$s_j = \sigma_j\left(\delta^*\right)$$

where $\delta^* \equiv (\delta_1^*, ..., \delta_J^*)$ solves the system of J independent equations

Endogeneity

- main problem: firms know ξ when they set prices so prices are correlated with the error which in turn is buried in a highly non-linear set of equations

- solution: assume M is large enough that

$$s_j = \sigma_j\left(\delta^*\right)$$

  where $\delta^* \equiv (\delta_1^*, ..., \delta_J^*)$ solves the system of J independent equations

- Berry shows that δ^* is unique and BLP provide a contraction mapping to find it

Estimation: F is known

- if we assume that F is known, then δ^*can be treated as a known nonlinear transformation of the market share data

Estimation: F is known

- if we assume that F is known, then δ^*can be treated as a known nonlinear transformation of the market share data

- using δ^* as data, run the regression

$$\delta^*(s) = x_j\beta - \alpha p_j + \xi_j$$

  - ξ is correlated with p, we need to find instruments: cost shifters, (exogenous) characteristics of other products

Estimation: General Case

- more generally, when F is not known, then δ^* depends on the unknown parameter λ of F

Estimation: General Case

- more generally, when F is not known, then δ^* depends on the unknown parameter λ of F
- for each value of (β, λ), there exists a unique solution of ξ that makes the predicted market shares equal to actual shares

Estimation: General Case

- more generally, when F is not known, then δ^* depends on the unknown parameter λ of F
- for each value of (β, λ), there exists a unique solution of ξ that makes the predicted market shares equal to actual shares
- let $\xi(\theta)$ denote this solution and then use the moment conditions

$$E[\xi(\theta_0) Z] = 0$$

  to estimate θ_0, where Z is a set of instrumental variables

Examples

- Logit model

$$\log(s_j) - \log(s_0) = \delta_j$$

  - no need to numerically compute δ's, simply run 2SLS of difference in log shares on (x_j, p_j) with instruments for price

Examples

- Logit model

$$\log(s_j) - \log(s_0) = \delta_j$$

  - no need to numerically compute δ's, simply run 2SLS of difference in log shares on (x_j, p_j) with instruments for price

- vertical model: $\delta_j = \varphi_j - p_j$
  - $s_j = F(\Delta_{j+1}) - F(\Delta_j)$ implies $\Delta_j = F^{-1}(F(\Delta_{j+1}) - s_j)$ with initial condition $\Delta_J = F^{-1}(1 - s_J)$
  - the values of φ_j can be obtained from the recursion $\varphi_j = \varphi_1 + \frac{p_j - p_{j-1}}{\Delta_j}$
  - treat φ_j as data and regress δ on x (use IVs if necessary)

Demand for Differentiated Products: Generation III Models

- our basic problem is that choices are correlated
  - if a consumer chooses product j, it means that she values the characteristics of product j, therefore, when p_j increases, she will tend to choose a product with similar characteristics

Demand for Differentiated Products: Generation III Models

- our basic problem is that choices are correlated
  - if a consumer chooses product j, it means that she values the characteristics of product j, therefore, when p_j increases, she will tend to choose a product with similar characteristics
- the correlation in individual choices occurs through the idiosyncratic term ε_{ij}: in the previous models, these shocks were assumed to be i.i.d
  - why not assume that they are correlated across products and estimate the variance-covariance matrix of ε_{ij}?

Demand for Differentiated Products: Generation III Models

- our basic problem is that choices are correlated
  - if a consumer chooses product j, it means that she values the characteristics of product j, therefore, when p_j increases, she will tend to choose a product with similar characteristics
- the correlation in individual choices occurs through the idiosyncratic term ε_{ij}: in the previous models, these shocks were assumed to be i.i.d
  - why not assume that they are correlated across products and estimate the variance-covariance matrix of ε_{ij}?
- the problem is that this approach simply reintroduces the dimensionality problem: have $\frac{J^2}{2}$ parameters to estimate

Random Coefficients

- an alternative approach is to impose structure on the form of correlation in idiosyncratic shocks: the idea is to specify the correlation as a linear function of the product characteristics

Random Coefficients

- an alternative approach is to impose structure on the form of correlation in idiosyncratic shocks: the idea is to specify the correlation as a linear function of the product characteristics
- consumer i's utility for product j

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

where $\beta_{ik} = \beta_k + \sigma_k \varsigma_{ik},\ k = 1, ..., K$

Random Coefficients

- an alternative approach is to impose structure on the form of correlation in idiosyncratic shocks: the idea is to specify the correlation as a linear function of the product characteristics

- consumer i's utility for product j

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

  where $\beta_{ik} = \beta_k + \sigma_k \varsigma_{ik},\ k = 1, ..., K$

- thus,

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

  where $\nu_{ij} = \sum_{k=1}^{K} x_{jk} \sigma_k \varsigma_{ik} + \epsilon_{ij}$
  - ϵ_{ij} is i.i.d. Type I extreme value, ς_{ik} is standard normal

Comments

- the idiosyncratic shock $\nu_{ij} = \sum_{k=1}^{K} x_{jk}\sigma_k\varsigma_{ik} + \epsilon_{ij}$ is correlated across products
  - if consumer i has a high realization of ς_{ik} for characteristic k, then she values this characteristic in all J products
  - consequently, if p_j increases, she will tend to switch to a product that has a lot of x_k

Comments

- the idiosyncratic shock $\nu_{ij} = \sum_{k=1}^{K} x_{jk}\sigma_k\varsigma_{ik} + \epsilon_{ij}$ is correlated across products
  - if consumer i has a high realization of ς_{ik} for characteristic k, then she values this characteristic in all J products
  - consequently, if p_j increases, she will tend to switch to a product that has a lot of x_k
- in modeling the correlation in this way, we have added K parameters to the model, one for each characteristic

Comments

- the idiosyncratic shock $\nu_{ij} = \sum_{k=1}^{K} x_{jk}\sigma_k \varsigma_{ik} + \epsilon_{ij}$ is correlated across products
  - if consumer i has a high realization of ς_{ik} for characteristic k, then she values this characteristic in all J products
  - consequently, if p_j increases, she will tend to switch to a product that has a lot of x_k

- in modeling the correlation in this way, we have added K parameters to the model, one for each characteristic

- the variation that identifies $\sigma = (\sigma_1, ..., \sigma_K)$ are changes in prices or products that generate substitution patterns that differ from those predicted by the logit model
  - if the data-generating model is logit, then we will estimate σ to be zero (i.e., the distributions of β_i is degenerate at β)

Estimation Algorithm

- integrate over ν_{ij} to obtain market shares

$$s_j(\delta, \theta) = \int \frac{\exp\left(\delta_j + \sum_{k=1}^{K} x_{jk}\sigma_k \varsigma_{ik}\right)}{1 + \sum_{m=1}^{J} \exp\left(\delta_m + \sum_{k=1}^{K} x_{mk}\sigma_k \varsigma_{ik}\right)} d\Phi(\varsigma)$$

  - this is K-dimension integral which cannot be computed easily

Estimation Algorithm

- integrate over ν_{ij} to obtain market shares

$$s_j(\delta,\theta) = \int \frac{\exp\left(\delta_j + \sum_{k=1}^{K} x_{jk}\sigma_k \varsigma_{ik}\right)}{1 + \sum_{m=1}^{J} \exp\left(\delta_m + \sum_{k=1}^{K} x_{mk}\sigma_k \varsigma_{ik}\right)} d\Phi(\varsigma)$$

  - this is K-dimension integral which cannot be computed easily

- Monte Carlo simulation: draw $\{\varsigma_{ik}^r\}_{r=1}^R$ from $\Phi(\cdot)$ and take the average

$$s_j(\delta,\theta) = \frac{1}{R}\sum_{r=1}^{R} \frac{\exp\left(\delta_j + \sum_{k=1}^{K} x_{jk}\sigma_k \varsigma_{ik}^r\right)}{1 + \sum_{m=1}^{J} \exp\left(\delta_m + \sum_{k=1}^{K} x_{mk}\sigma_k \varsigma_{ik}^r\right)}$$

Estimation Algorithm

- integrate over ν_{ij} to obtain market shares

$$s_j(\delta, \theta) = \int \frac{\exp\left(\delta_j + \sum_{k=1}^{K} x_{jk}\sigma_k \varsigma_{ik}\right)}{1 + \sum_{m=1}^{J} \exp\left(\delta_m + \sum_{k=1}^{K} x_{mk}\sigma_k \varsigma_{ik}\right)} d\Phi(\varsigma)$$

  - this is K-dimension integral which cannot be computed easily

- Monte Carlo simulation: draw $\{\varsigma_{ik}^r\}_{r=1}^{R}$ from $\Phi(\cdot)$ and take the average

$$s_j(\delta, \theta) = \frac{1}{R} \sum_{r=1}^{R} \frac{\exp\left(\delta_j + \sum_{k=1}^{K} x_{jk}\sigma_k \varsigma_{ik}^r\right)}{1 + \sum_{m=1}^{J} \exp\left(\delta_m + \sum_{k=1}^{K} x_{mk}\sigma_k \varsigma_{ik}^r\right)}$$

- equate actual to simulated market shares and invert the system to obtain the mean utilities, or equivalently $\xi(\theta, s)$, then interact $\xi(\theta, s)$ with instruments z and find the value of θ that makes the sample moments as close to 0 as possible

Remarks

- the simulated market shares enter non-linearly in the moment conditions so nice properties of simulation estimators are not valid

Remarks

- the simulated market shares enter non-linearly in the moment conditions so nice properties of simulation estimators are not valid
- IIA property continues to hold at the individual level: ratio of choice probabilities does not depend upon number or utility of the other alternatives

Remarks

- the simulated market shares enter non-linearly in the moment conditions so nice properties of simulation estimators are not valid
- IIA property continues to hold at the individual level: ratio of choice probabilities does not depend upon number or utility of the other alternatives
- but, market shares no longer have the IIA property, aggregating over the realizations of ς implies that ratio of market shares depends upon the number and characteristics of alternative products

Data on Consumer Characteristics

- in many cases, we can observe the distribution of consumer characteristics
  - e.g., from census data, we can learn the income, education, age, and family size of consumers living in a particular geographical market

Data on Consumer Characteristics

- in many cases, we can observe the distribution of consumer characteristics
  - e.g., from census data, we can learn the income, education, age, and family size of consumers living in a particular geographical market

- let z_i denote the vector of observable consumer characteristics, then our model of how consumer preferences vary as a function of observed and unobserved individual characteristics is that

$$\nu_{ij} = \sum_{k=1}^{K} x_{jk} \left(\pi_k z_{ik} + \sigma_k \varsigma_{ik} \right) + \epsilon_{ij}$$

Data on Consumer Characteristics

- in many cases, we can observe the distribution of consumer characteristics
  - e.g., from census data, we can learn the income, education, age, and family size of consumers living in a particular geographical market

- let z_i denote the vector of observable consumer characteristics, then our model of how consumer preferences vary as a function of observed and unobserved individual characteristics is that

$$\nu_{ij} = \sum_{k=1}^{K} x_{jk} \left(\pi_k z_{ik} + \sigma_k \varsigma_{ik} \right) + \epsilon_{ij}$$

- the choice probabilities for consumer i are obtained by integrating over the idiosyncratic shock ϵ as above

Estimation Algorithm

- to obtain the market share of product j, we need to integrate over
  - the unobserved characteristics ς which are distributed as standard normal
  - the observed characteristics which are distributed in the population according to some joint distribution G, this distribution is obtained (parametrically or non-parametrically) from census data

Estimation Algorithm

- to obtain the market share of product j, we need to integrate over
  - the unobserved characteristics ς which are distributed as standard normal
  - the observed characteristics which are distributed in the population according to some joint distribution G, this distribution is obtained (parametrically or non-parametrically) from census data
- estimation
  - draw vectors of consumer characteristics from these distributions, determine individual choices
  - aggregate to obtain predicted market shares
  - solve demand system to obtain $\xi(\theta, s)$ and then interact with instruments (x, w) to do GMM

Remarks

- the demographic information reduces the reliance on parametric assumptions about the distribution of consumer heterogeneity

Remarks

- the demographic information reduces the reliance on parametric assumptions about the distribution of consumer heterogeneity
- it also allows the model to incorporate differences in the distribution of consumers across markets and their impact on aggregate demand
  - for example, all empirical evidence suggests that the impact of price on consumer demand depends on the consumer's income
  - so if the distribution of income varies across geographical market, then each market has a different price coefficient
  - the random coefficients model with demographic characteristics captures this interaction

Remarks (Cont.)

- it provides an approximation to the demand surface that is tailored to each market and does not impose one approximation to all markets
  - better fit leads to more precise parameter estimates
  - provides a tool for making predictions of likely outcomes in new markets or from policies that would affect the distribution of consumer characteristics

Pricing Equations

- suppose there are N firms in the market, indexed by t
  - firms may produce more than one product, let J_t denote the number of products by firm t
  - firms choose prices, let p_t denote the price vector for firm t and p_{-t} the prices of its rivals

Pricing Equations

- suppose there are N firms in the market, indexed by t
  - firms may produce more than one product, let J_t denote the number of products by firm t
  - firms choose prices, let p_t denote the price vector for firm t and p_{-t} the prices of its rivals

- each firm t choose p_t to maximize

$$\pi_t(p_t, p_{-t}) = \sum_{j \in J_t} p_j M s_j(x, p, \xi) - C_t(q_t, x_t)$$

FOC

- first-order equations for product j is

$$s_j\left(x, p, \xi\right) + \sum_{r \in J_t} \left(p_r - mc_r\right) \frac{\partial s_r}{\partial p_j} = 0$$

FOC

- first-order equations for product j is

$$s_j\left(x, p, \xi\right) + \sum_{r \in J_t} \left(p_r - mc_r\right) \frac{\partial s_r}{\partial p_j} = 0$$

- in matrix notation

$$s + \left(p - mc\right) \Delta = 0$$

where Δ_{ij} is nonzero for the elements of a row that are produced by the same firm as the row good (diagonal if each firm produces only one good)

Estimation Equation

- to complete the model, we need to specify a functional form for marginal cost

$$\log(mc_j) = w_j \gamma + \omega_j$$

Estimation Equation

- to complete the model, we need to specify a functional form for marginal cost

$$\log (mc_j) = w_j\gamma + \omega_j$$

- substituting back, we obtain the pricing equations

$$\log \left(p - \Delta^{-1}s\right) - w\gamma = \omega (\theta)$$

Estimation Equation

- to complete the model, we need to specify a functional form for marginal cost

$$\log(mc_j) = w_j\gamma + \omega_j$$

- substituting back, we obtain the pricing equations

$$\log\left(p - \Delta^{-1}s\right) - w\gamma = \omega(\theta)$$

- note that Δ is the derivative of market demand so it depends on the demand parameters
  - the pricing and demand equations can be estimated jointly using simulated method of moment estimator

Remarks

- in practice, the inclusion of the pricing equations really helps identify the demand parameters of the random coefficient model
  - the demand model is often too flexible for the data: not enough variation across products and markets relative to the approximations

Remarks

- in practice, the inclusion of the pricing equations really helps identify the demand parameters of the random coefficient model
  - the demand model is often too flexible for the data: not enough variation across products and markets relative to the approximations
- in some cases, the authors does not estimate product marginal costs but back them out estimates from FOC directly

BLP (1995)

- goal: provide a framework for obtaining estimates of demand and cost parameters for a class of oligopolistic differentiated good markets using only aggregate data on product shares and prices

BLP (1995)

- goal: provide a framework for obtaining estimates of demand and cost parameters for a class of oligopolistic differentiated good markets using only aggregate data on product shares and prices
- extends the literature in two important ways
  - relax the strong functional form assumptions that restrict the substitution pattern
  - accounts for the endogeneity of prices

Data

- product characteristics: number of cylinders, number of doors, weight, engine displacement, horsepower, length, width, wheelbase, EPA miles per gallon rating, and indicator variables for whether the car has front wheel drive, automatic transmission, power steering and air conditioning

Data

- product characteristics: number of cylinders, number of doors, weight, engine displacement, horsepower, length, width, wheelbase, EPA miles per gallon rating, and indicator variables for whether the car has front wheel drive, automatic transmission, power steering and air conditioning
- price is list retail price for the base model, measured in 1983 dollar, quantity is sales in the US

Data

- product characteristics: number of cylinders, number of doors, weight, engine displacement, horsepower, length, width, wheelbase, EPA miles per gallon rating, and indicator variables for whether the car has front wheel drive, automatic transmission, power steering and air conditioning

- price is list retail price for the base model, measured in 1983 dollar, quantity is sales in the US

- all models marketed during the 20 year period from 1971 to 1990, treating model/year as an observation, total sample size is 2217, number of distinct models is 997

Data

- product characteristics: number of cylinders, number of doors, weight, engine displacement, horsepower, length, width, wheelbase, EPA miles per gallon rating, and indicator variables for whether the car has front wheel drive, automatic transmission, power steering and air conditioning
- price is list retail price for the base model, measured in 1983 dollar, quantity is sales in the US
- all models marketed during the 20 year period from 1971 to 1990, treating model/year as an observation, total sample size is 2217, number of distinct models is 997
- additional data: price of gasoline, number of HH in US, etc.

Data Overview

DESCRIPTIVE STATISTICS

Year	No. of Models	Quantity	Price	Domestic	Japan	European	HP/Wt	Size	Air	MPG	MP$
1971	92	86.892	7.868	0.866	0.057	0.077	0.490	1.496	0.000	1.662	1.850
1972	89	91.763	7.979	0.892	0.042	0.066	0.391	1.510	0.014	1.619	1.875
1973	86	92.785	7.535	0.932	0.040	0.028	0.364	1.529	0.022	1.589	1.819
1974	72	105.119	7.506	0.887	0.050	0.064	0.347	1.510	0.026	1.568	1.453
1975	93	84.775	7.821	0.853	0.083	0.064	0.337	1.479	0.054	1.584	1.503
1976	99	93.382	7.787	0.876	0.081	0.043	0.338	1.508	0.059	1.759	1.696
1977	95	97.727	7.651	0.837	0.112	0.051	0.340	1.467	0.032	1.947	1.835
1978	95	99.444	7.645	0.855	0.107	0.039	0.346	1.405	0.034	1.982	1.929
1979	102	82.742	7.599	0.803	0.158	0.038	0.348	1.343	0.047	2.061	1.657
1980	103	71.567	7.718	0.773	0.191	0.036	0.350	1.296	0.078	2.215	1.466
1981	116	62.030	8.349	0.741	0.213	0.046	0.349	1.286	0.094	2.363	1.559
1982	110	61.893	8.831	0.714	0.235	0.051	0.347	1.277	0.134	2.440	1.817
1983	115	67.878	8.821	0.734	0.215	0.051	0.351	1.276	0.126	2.601	2.087
1984	113	85.933	8.870	0.783	0.179	0.038	0.361	1.293	0.129	2.469	2.117
1985	136	78.143	8.938	0.761	0.191	0.048	0.372	1.265	0.140	2.261	2.024
1986	130	83.756	9.382	0.733	0.216	0.050	0.379	1.249	0.176	2.416	2.856
1987	143	67.667	9.965	0.702	0.245	0.052	0.395	1.246	0.229	2.327	2.789
1988	150	67.078	10.069	0.717	0.237	0.045	0.396	1.251	0.237	2.334	2.919
1989	147	62.914	10.321	0.690	0.261	0.049	0.406	1.259	0.289	2.310	2.806
1990	131	66.377	10.337	0.682	0.276	0.043	0.419	1.270	0.308	2.270	2.852
All	2217	78.804	8.604	0.790	0.161	0.049	0.372	1.357	0.116	2.099	2.086

- number of products rises from 72 in 1974 to high of 150 in 1988, sales per model trend down

Data Overview

- list prices have risen almost 50 percent during the 1980s but characteristics are also changing so not clear what is happening to cost per car with fixed characteristics
- HP/weight trended down and then up, mostly due to changes in weight, fuel efficiency trends up
- air conditioning is increasingly part of the base model
- market share of domestics fall from 93% to 68%, mostly to Japanese models

OLS and IV Logit Results

RESULTS WITH LOGIT DEMAND AND MARGINAL COST PRICING
(2217 OBSERVATIONS)

Variable	OLS Logit Demand	IV Logit Demand	OLS ln(*price*) on *w*
Constant	−10.068	−9.273	1.882
	(0.253)	(0.493)	(0.119)
*HP/Weight**	−0.121	1.965	0.520
	(0.277)	(0.909)	(0.035)
Air	−0.035	1.289	0.680
	(0.073)	(0.248)	(0.019)
MP$	0.263	0.052	—
	(0.043)	(0.086)	
*MPG**	—	—	−0.471
			(0.049)
*Size**	2.341	2.355	0.125
	(0.125)	(0.247)	(0.063)
Trend	—	—	0.013
			(0.002)
Price	−0.089	−0.216	—
	(0.004)	(0.123)	
No. Inelastic Demands	1494	22	*n.a.*
(+/−2 *s.e.'s*)	(1429−1617)	(7−101)	
R^2	0.387	*n.a.*	.656

OLS and IV Logit Results

- OLS estimates
  - most of the estimates have the right sign but not very precisely estimated
  - price coefficient is implausibly small: 1494 of the 2217 models have inelastic demands, which is not consistent with profit-maximization
  - 61 percent of the variance in mean utility is due to unobserved product characteristics

OLS and IV Logit Results

- OLS estimates
  - most of the estimates have the right sign but not very precisely estimated
  - price coefficient is implausibly small: 1494 of the 2217 models have inelastic demands, which is not consistent with profit-maximization
  - 61 percent of the variance in mean utility is due to unobserved product characteristics
- IV results
  - all characteristics enter positively and significantly (except for MP$)
  - price coefficient increases: products with higher unobserved quality sell for higher prices
  - number of products with inelastic demands drops to 22

Random Coefficient Model with Pricing Equations

Estimated Parameters of the Demand and Pricing Equations:
BLP Specification, 2217 Observations

Demand Side Parameters	Variable	Parameter Estimate	Standard Error	Parameter Estimate	Standard Error
Means ($\bar{\beta}$'s)	*Constant*	−7.061	0.941	−7.304	0.746
	HP/Weight	2.883	2.019	2.185	0.896
	Air	1.521	0.891	0.579	0.632
	MP$	−0.122	0.320	−0.049	0.164
	Size	3.460	0.610	2.604	0.285
Std. Deviations (σ_β's)	*Constant*	3.612	1.485	2.009	1.017
	HP/Weight	4.628	1.885	1.586	1.186
	Air	1.818	1.695	1.215	1.149
	MP$	1.050	0.272	0.670	0.168
	Size	2.056	0.585	1.510	0.297
Term on Price (α)	$\ln(y-p)$	43.501	6.427	23.710	4.079
Cost Side Parameters					
	Constant	0.952	0.194	0.726	0.285
	ln(*HP/Weight*)	0.477	0.056	0.313	0.071
	Air	0.619	0.038	0.290	0.052
	ln(*MPG*)	−0.415	0.055	0.293	0.091
	ln(*Size*)	−0.046	0.081	1.499	0.139
	Trend	0.019	0.002	0.026	0.004
	$\ln(q)$			−0.387	0.029

- the standard deviations of the random coefficients are quite important

Substitution Patterns

A SAMPLE FROM 1990 OF ESTIMATED OWN- AND CROSS-PRICE SEMI-ELASTICITIES:
BASED ON TABLE IV (CRTS) ESTIMATES

	Mazda 323	Nissan Sentra	Ford Escort	Chevy Cavalier	Honda Accord	Ford Taurus	Buick Century	Nissan Maxima	Acura Legend	Lincoln Town Car	Cadillac Seville	Lexus LS400	BMW 735i
323	−125.933	1.518	8.954	9.680	2.185	0.852	0.485	0.056	0.009	0.012	0.002	0.002	0.000
Sentra	0.705	−115.319	8.024	8.435	2.473	0.909	0.516	0.093	0.015	0.019	0.003	0.003	0.000
Escort	0.713	1.375	−106.497	7.570	2.298	0.708	0.445	0.082	0.015	0.015	0.003	0.003	0.000
Cavalier	0.754	1.414	7.406	−110.972	2.291	1.083	0.646	0.087	0.015	0.023	0.004	0.003	0.000
Accord	0.120	0.293	1.590	1.621	−51.637	1.532	0.463	0.310	0.095	0.169	0.034	0.030	0.005
Taurus	0.063	0.144	0.653	1.020	2.041	−43.634	0.335	0.245	0.091	0.291	0.045	0.024	0.006
Century	0.099	0.228	1.146	1.700	1.722	0.937	−66.635	0.773	0.152	0.278	0.039	0.029	0.005
Maxima	0.013	0.046	0.236	0.256	1.293	0.768	0.866	−35.378	0.271	0.579	0.116	0.115	0.020
Legend	0.004	0.014	0.083	0.084	0.736	0.532	0.318	0.506	−21.820	0.775	0.183	0.210	0.043
TownCar	0.002	0.006	0.029	0.046	0.475	0.614	0.210	0.389	0.280	−20.175	0.226	0.168	0.048
Seville	0.001	0.005	0.026	0.035	0.425	0.420	0.131	0.351	0.296	1.011	−16.313	0.263	0.068
LS400	0.001	0.003	0.018	0.019	0.302	0.185	0.079	0.280	0.274	0.606	0.212	−11.199	0.086
735*i*	0.000	0.002	0.009	0.012	0.203	0.176	0.050	0.190	0.223	0.685	0.215	0.336	−9.376

Note: Cell entries i, j, where i indexes row and j column, give the percentage change in market share of i with a \$1000 change in the price of j.

Remarks

- cross-price elasticities are large for cars with similar characteristics
- magnitudes of the impact of price increases of the higher price cars are much smaller than they are for the lower-priced cars
- patterns seem plausible: Lexus is closest substitute for BMW 735, Accord is the closest substitute for Taurus

Markups

A SAMPLE FROM 1990 OF ESTIMATED PRICE-MARGINAL COST MARKUPS AND VARIABLE PROFITS: BASED ON TABLE 6 (CRTS) ESTIMATES

	Price	Markup Over MC $(p - MC)$	Variable Profits (in $'000's) $q * (p - MC)$
Mazda 323	$5,049	$ 801	$18,407
Nissan Sentra	$5,661	$ 880	$43,554
Ford Escort	$5,663	$1,077	$311,068
Chevy Cavalier	$5,797	$1,302	$384,263
Honda Accord	$9,292	$1,992	$830,842
Ford Taurus	$9,671	$2,577	$807,212
Buick Century	$10,138	$2,420	$271,446
Nissan Maxima	$13,695	$2,881	$288,291
Acura Legend	$18,944	$4,671	$250,695
Lincoln Town Car	$21,412	$5,596	$832,082
Cadillac Seville	$24,353	$7,500	$249,195
Lexus LS400	$27,544	$9,030	$371,123
BMW 735*i*	$37,490	$10,975	$114,802

Remarks and Conclusions

- remarks on the markup
  - average markup is \$3,753 and average ratio of markup to retail price is .239
  - patterns are plausible: markups are higher on higher-priced models

Remarks and Conclusions

- remarks on the markup
  - average markup is $3,753 and average ratio of markup to retail price is .239
  - patterns are plausible: markups are higher on higher-priced models
- conclusions
  - price endogeneity matters
  - allowing for more flexible utility specifications generates a more realistic picture of demand and equilibrium

Reading for Next Class

- A. Nevo and C. Wolfram, "Why Do Manufacturers Issue Coupons? An Empirical Analysis of Breakfast Cereals," *RAND Journal of Economics*, 2002.