Empirical Industrial Organization

Pierre Dubois

Toulouse School of Economics

Spring 2018

Course Outline

- **Part I. Demand for differentiated products and IO applications**

  1. **Differentiated products demand**
     1.1. **Theory and estimation on micro data**
     1.2. **Theory and estimation on aggregate data**
  2. **Measuring market power and merger analysis**
     2.1. **Market power estimation**
     2.2. **Merger analysis and simulation**

Course Outline

3. Measuring consumer welfare

4. Identifying contracts in vertical relations

5. Consumer demand with limited information and advertising

6. Applications on Industry and Trade

7. Discrete/continuous Demand Models

8. Dynamic Demand

Introduction: why is Demand Useful in IO

- Allows to "reverse engineer" firms' optimal decisions in order to
  - obtain marginal costs
  - test models of pricing
- Compute firms strategies that depend on consumer behavior
  - price discrimination
  - advertising and promotional activity
- Simulate counterfactuals
  - likely effect of mergers
  - demand for new products
  - government regulation
- Compute consumer welfare

Background

Most straightforward approach is to specify a demand function directly:

$$q = D(p, r, \varepsilon)$$

- q is a $J \times 1$ vector of quantities
- p is a $J \times 1$ vector of prices
- r is a vector of exogenous variables
- ε is a $J \times 1$ vector of random shocks.

Early work focused on how to specify $D(\cdot)$ in a way that was both flexible and consistent with economic theory.

Linear Expenditure model (Stone, 1954), Rotterdam model (Theil, 1965; Barten 1966), Translog model (Christensen, Jorgenson, and Lau, 1975), Almost Ideal Demand System (Deaton and Muellbauer, 1980), ...

Issues

Issues for many cases considered in IO:

"The too many parameters problem". As the number of options, J, becomes large there are too many parameters to estimate

- Even if $D(p, r, \varepsilon) = Ap + \varepsilon$, where A is $J \times J$ matrix of parameters, implies J^2 parameters to be estimated.
- With a more flexible functional form, the problem is even greater.

Heterogeneity. for some applications we would like to explicitly model and estimate the distribution of heterogeneity.

- Above approach, generally, does not let us do this
- Presence of heterogeneity doesn't invalidate aggregate demand but
  - should care in imposing the restrictions of economic theory
  - for example if preferences are of Gorman form (Gorman, 1959)

Issues

Explicit modeling of specific consumer behavior. easier when we start with an explicit model of consumer behavior and aggregate to market level demand.

New Goods. This demand system does not easily allow us to predict the demand for new goods.

Estimation. need to include highly colinear prices, and also instrument for them

Approaches

- Modelling of demand either in *product space* or in *characteristics space*
- Implicit vs. explicit modeling of heterogeneity
- Modelling choices and trade-offs to be addressed:
    - aggregation across products
    - symmetry across products
    - weak separability across products and multi stage budgeting
    - models in characteristics space and discrete choice

Modelling Approach in IO

- Random utility model underlying decisions of demand
- Dimensionality reduction with characteristics approach
- Usually implemented as discrete choice, but does not have to be
- Consumer choice will depend on the best alternative among a given choice set

Models in Characteristics Space

- Previous examples are models in product space
- Models in characteristics space: the utility function is a function of the attributes of the product (Gorman, 1956/1980, Lancaster, 1966)
    - Connections to hedonics
    - Usually implemented as discrete choice, but does not have to be
- If the product identity is a characteristic, then approach nests the product space approach
- Indirect utility model at the level of the decision maker
- Modelling in products spaces using characteristics space for flexibility (Dubois, Griffith, Nevo, 2014)

Models in Characteristics Space

- Indirect utility function of consumer i, from product j in market t:

$$U(x_{jt}, \xi_{jt}, I_i - p_{jt}, \tau_i; \theta)$$

where

x_{jt} – $1 \times K$ vector of observed product characteristics
ξ_{jt} – unobserved (by us) product characteristic
τ_i – individual characteristics
I_i – income

- ξ_{jt} will play an important role
  - realistic (inability of observed characteristics to capture the essence of the product)
  - will act as a "residual" (why predicted shares do not fit exactly)
  - potentially implies endogeneity

Choice and Normalizations

- Consumer chooses j if

$$U(x_{jt}, \xi_{jt}, I_i - p_{jt}, \tau_i; \theta) \geq U(x_{kt}, \xi_{kt}, I_i - p_{kt}, \tau_i; \theta)$$

  for all $k \neq j$.
- For what follows utility is ordinal and not cardinal, and therefore is invariant to affine transformation
- This means that we typically have 2 normalizations
- In what follows (will make more sense below)
  - normalize the variance of one component to one
  - normalize the utility from the outside good to zero

Logit Model

- A common model of this class is the Logit model

Individual random utility for the $J+1$ goods can be written ($j = 0, ..J$)

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

where ε_{ij} is a stochastic term
(equivalent to $u_{ij} = x_j \beta - \alpha (y_i - p_j) + \xi_j + \varepsilon_{ij}$ because all variables non specific to alternative cancel out)

- 2 views of ε_{ij}:
    - utility is deterministic, but the choice process itself is probabilistic (Tversky, 1972)
    - utility is deterministic, but ε_{ij} captures the researcher's inability to formulate individual behavior precisely
- Interplay between ξ_j and ε_{ij}: in a way all what ξ_j is doing is changing the mean of ε_{ij}, by j.

Logit Model

- Assume individual chooses the product that maximizes random utility and derive individual choice probabilities
- Aggregate market shares can be derived from individual choice probabilities
- Here, all consumers have same valuations for the characteristics and price, we can write

$$u_{ij} = \delta_j + \varepsilon_{ij}$$

  where $\delta_j = x_j\beta - \alpha p_j + \xi_j$ is the mean or common part of consumers' utility.
- The mean utility of the outside good is normalized to 0 so $\delta_0 = 0$.
- In the logit model, consumers only differ in ε_{ij}.

Logit Model

- ε_{ij} is modeled as an i.i.d. random variable with an extreme value distribution

$$F(\varepsilon_{ij}) = \exp(-\exp(-\varepsilon_{ij}))$$

- Choice probability

$$\begin{aligned} P_{ij} &= P\left(u_{ij} \geq u_{ik}, \forall k \neq j\right) \\ &= P\left(\varepsilon_{ik} - \varepsilon_{ij} \leq \delta_j - \delta_k, \forall k \neq j\right) \end{aligned}$$

- The difference between two extreme value variables is distributed logistic such that

$$F(\varepsilon_{ij} - \varepsilon_{ik}) = \frac{\exp\left(\varepsilon_{ij} - \varepsilon_{ik}\right)}{1 + \exp\left(\varepsilon_{ij} - \varepsilon_{ik}\right)}$$

Logit Model

- Logit choice probabilities

$$
\begin{aligned}
P_{ij} &= P\left(u_{ij} \geq u_{ik}, \forall k \neq j\right) = P\left(\varepsilon_{ik} - \varepsilon_{ij} \leq \delta_j - \delta_k, \forall k \neq j\right) \\
&= P\left(\varepsilon_{ik} \leq \varepsilon_{ij} + \delta_j - \delta_k, \forall k \neq j\right) \\
&= \int P\left(\varepsilon_{ik} \leq \varepsilon_{ij} + \delta_j - \delta_k, \forall k \neq j | \varepsilon_{ij}\right) dF\left(\varepsilon_{ij}\right) \\
&= \int \prod_{k \neq j} \exp(-\exp(-\varepsilon_{ij} - \delta_j + \delta_k)) dF\left(\varepsilon_{ij}\right) \\
&= \int \prod_{k \neq j} \exp(\text{-}\exp(\text{-}\varepsilon_{ij}\text{-}\delta_j + \delta_k)) \exp(\text{-}\exp(\text{-}\varepsilon_{ij})) \exp(\text{-}\varepsilon_{ij}) d\varepsilon_{ij}
\end{aligned}
$$

- Details in book of Train (2009), page 74

Logit Model

$$
\begin{aligned}
P_{ij} &= \int_{-\infty}^{+\infty} \left(\prod_{k \neq j} \exp(-\exp(-\varepsilon_{ij}-\delta_j+\delta_k)) \right) \exp(-\exp(-\varepsilon_{ij})) \exp(-\varepsilon_{ij}) d\varepsilon_{ij} \\
&= \int_{-\infty}^{+\infty} \exp(-\sum_k \exp(-\varepsilon_{ij}-\delta_j+\delta_k)) \exp(-\varepsilon_{ij}) d\varepsilon_{ij} \\
&= \int_{-\infty}^{+\infty} \exp(-\exp(-\varepsilon_{ij}) \sum_k \exp(-\delta_j+\delta_k)) \exp(-\varepsilon_{ij}) d\varepsilon_{ij} \\
&= \int_{0}^{+\infty} \exp(-\theta \sum_k \exp(-\delta_j+\delta_k)) d\theta \text{ (change of variable)} \\
&= \left[\frac{\exp(-\theta \sum_k \exp(-\delta_j+\delta_k))}{-\sum_k \exp(-\delta_j+\delta_k)} \right]_0^{+\infty} = \frac{1}{\sum_{k=0}^{J} \exp(\delta_k - \delta_j)} \\
P_{ij} &= \frac{\exp \delta_j}{1+\sum_{k=1}^{J} \exp \delta_k} \in (0,1)
\end{aligned}
$$

Logit Model

- IIA property

$$\frac{P_{ij}}{P_{ij'}} = \exp\left(\delta_j - \delta_{j'}\right)$$

- Consumer surplus

$$CS = \frac{1}{\alpha} \max_j u_{ij}$$

where α is the marginal utility of income

- Then

$$\begin{aligned} E\left[CS\right] &= \frac{1}{\alpha} E\left[\max_j u_{ij}\right] \\ &= \frac{1}{\alpha} E\left[\max_j \left(\delta_j + \varepsilon_{ij}\right)\right] \\ &= \frac{1}{\alpha} \ln \sum_{k=1}^{J} \exp \delta_k + C \end{aligned}$$

Logit Model

- Change in consumer surplus between 0 and 1

$$\Delta E\left[CS\right] = \frac{1}{\alpha}\left(\ln\sum_{k=1}^{J^1}\exp\delta_k^1 - \ln\sum_{k=1}^{J^0}\exp\delta_k^0\right)$$

Logit Model

- Choice probabilities

$$P_{ij}\left(\alpha,\beta,\xi\right)=\frac{\exp\left(x_j\beta-\alpha p_j+\xi_j\right)}{1+\sum_{k=1}^{J}\exp\left(x_k\beta-\alpha p_k+\xi_k\right)}$$

and for the outside good

$$P_{i0}\left(\alpha,\beta,\xi\right)=\frac{1}{1+\sum_{k=1}^{J}\exp\left(x_k\beta-\alpha p_k+\xi_k\right)}$$

- McFadden (1974) proved that the log-likelihood function with these choice probabilities is globally concave in parameters (α,β,ξ).

Logit Model Estimation

- Estimation by Maximum Likelihood with N observations (decision makers) is possible if sample is exogenously drawn

$$\max_{\alpha,\beta,\xi} \sum_{i=1}^{N} \sum_{j} d_{ij} \ln P_{ij}\left(\alpha, \beta, \xi\right)$$

where $d_{ij} = 1$ if i chooses j and 0 otherwise

Logit Model Estimation: Aggregate Data

- The aggregate market share function is simply equal to the individual choice probability, as a function of the vector δ.
- With L consumers, the observed market share (relative to the potential market) is simply

$$s_j = \frac{q_j}{L}$$

- Equating market share functions to observed market shares (in vector form)

$$P(\alpha, \beta, \xi) = s$$

but error terms ξ_j enter non linearly through the mean utilities δ.

Logit Model Estimation: Aggregate Data

- Berry (1994) showed how to invert the demand system easily to solve for the mean utility vector
- We have

$$s_j = \frac{\exp\left(x_j\beta - \alpha p_j + \xi_j\right)}{1 + \sum_{k=1}^{J} \exp\left(x_k\beta - \alpha p_k + \xi_k\right)}$$

$$s_0 = \frac{1}{1 + \sum_{k=1}^{J} \exp\left(x_k\beta - \alpha p_k + \xi_k\right)}$$

- So that

$$\frac{s_j}{s_0} = \exp\left(x_j\beta - \alpha p_j + \xi_j\right)$$

$$\ln\left(\frac{s_j}{s_0}\right) = x_j\beta - \alpha p_j + \xi_j$$

- We obtain a standard linear regression model

Logit Model Estimation: Aggregate Data

- We observe quantities demanded and not market shares, but if we know the market size L then

$$\ln\left(\frac{q_j / L}{q_0 / L}\right) = \ln\left(\frac{q_j}{L - \sum_{k=1}^{J} q_k}\right)$$

- Estimation with different markets (or periods):

$$\ln\left(\frac{s_{jt}}{s_{0t}}\right) = x_{jt}\beta - \alpha p_{jt} + \xi_{jt}$$

- Time (market) dummies allow to capture unknown market size.

Logit Model Estimation: Aggregate Data

- Linear equation

$$\ln\left(\frac{s_{jt}}{s_{0t}}\right) = x_{jt}\beta - \alpha p_{jt} + \xi_{jt}$$

  The price may be correlated with error term, such that OLS would lead to an estimate of α which is biased towards zero.
- Solution: 2SLS
- Instruments to identify the parameters:
  - cost side variables (not always available with product level data)
  - characteristics of competitors (see BLP 1995): x_j and $\sum_{k\neq j} x_k$ and J and similar sums by firm
  - lagged variables or variables from other markets (if panel data)

Logit Model Estimation: Aggregate Data

- As
$$\frac{\partial s_j}{\partial p_j} = -\alpha s_j \left(1 - s_j\right) \text{ and } \frac{\partial s_k}{\partial p_j} = \alpha s_j s_k$$
- Elasticities have simple expressions
$$\eta_{jk} = \frac{\partial s_j}{\partial p_k} \frac{p_k}{s_j} = \begin{cases} -\alpha p_j (1 - s_j) & \text{if } j = k \\ \alpha p_k s_k & \text{otherwise} \end{cases}$$
- Problems:
  - Own price elasticities: market shares are small, so $\alpha(1 - s_j)$ is nearly constant and therefore the own-price elasticities are proportional to price. This is driven mostly by lack of heterogeneity
  - Cross-price elasticities: cross price elasticity w.r.t. a change in the price of product k is that same for all products such that $j \neq k$. This is driven by lack of heterogeneity and i.i.d.

Nested Logit Model

- One way to relax Logit assumptions with model still easy to estimate is the Nested Logit Model
- The individual specific error ε_{ij} now follows a more general extreme value distribution
- Example with one level of nests, groups $g = 0, .., G$
- Substitution patterns:
  - for alternatives in same nest, ratio of probabilities independent of other alternatives (IIA within nest)
  - for alternatives in different nests, ratio of probabilities can depend on attributes of other alternatives in the two nests

Nested Logit Model

- Group 0 is for the outside good, with one product only
- The c.d.f. of $(\varepsilon_{i0}, .., \varepsilon_{ij}, .., \varepsilon_{iJ})$ is (GEV)

$$F(\varepsilon_{i0}, .., \varepsilon_{ij}, .., \varepsilon_{iJ}) = \exp\left(-\sum_{g=0}^{G}\left(\sum_{j\in B_g}\exp\left(-\frac{\varepsilon_{ij}}{\lambda_g}\right)\right)^{\lambda_g}\right)$$

- $\lambda_g \in [0, 1]$ measures the degree of independence in unobserved utility among the alternatives in nest g
- One could always reinterpret ε using Cardell (1997) showing that we can write $\varepsilon_{ij} = \nu_{ij} + \lambda_g \eta_{ig(j)}$
- Cardell (1997):
  - For $\lambda \in (0, 1)$ and η extreme value type I there exists a unique distribution or a random variable ν independent of η, such that $\nu + \lambda\eta$ is type I extreme value.
  - The p.d.f. of ν is $f_\lambda(\nu) = \frac{1}{\lambda}\sum_{n=0}^{\infty}\frac{(-1)^n \exp(-n\nu)}{n!\Gamma(-\lambda n)}$

Nested Logit Model

- Choice probabilities:

$$P_{ij}\left(\alpha,\beta,\xi\right)=\frac{\exp\left(\delta_j/\lambda_g\right)\left(\sum_{k\in B_g}\exp\left(\frac{\delta_k}{\lambda_g}\right)\right)^{\lambda_g-1}}{\sum_{l=1}^{G}\left(\sum_{k\in B_l}\exp\left(\frac{\delta_k}{\lambda_l}\right)\right)^{\lambda_l-1}}$$

- Decomposition in two Logits is possible

Nested Logit Model

- Decompose δ_j into ω_g fixed within nest and θ_j which varies within nest: $\delta_j = \omega_g + \theta_j$
- Using

$$P_{ij} = P_{ij|B_g} P_{iB_g}$$

- We can show that

$$P_{iB_g} = \frac{\exp\left(\omega_g + \lambda_g I_g\right)}{\sum_{l=1}^{G} \exp\left(\omega_l + \lambda_l I_l\right)}$$

where $I_g = \ln \sum_{k \in B_g} \exp\left(\frac{\theta_k}{\lambda_g}\right)$ is the "inclusive value" of nest g

- And

$$P_{ij|B_g} = \frac{\exp\left(\theta_j / \lambda_g\right)}{\sum_{k \in B_g} \exp\left(\frac{\theta_k}{\lambda_g}\right)}$$

- Overlapping nests is possible

Nested Logit Model Estimation: Aggregate Data

- Estimation by Maximum Likelihood on individual data is possible
- We can also invert the system of aggregate market shares and solve for δ.
- Berry (1994) showed that we have the following analytical solution

$$\ln\left(\frac{s_j}{s_0}\right) = x_j\beta - \alpha p_j + \sigma_g \ln s_{j|g} + \xi_j$$

where $s_{j|g} = s_j / s_g$ and $\sigma_g = 1 - \lambda_g$

- We obtain a linear regression but like in Logit OLS are biased
- Both p_j and $\ln s_{j|g}$ are endogenous
- Estimation by 2SLS possible
- Need additional instruments: for example, sum of competitors characteristics by groups

Nested Logit Model : Aggregate Data

- Own and cross price elasticities have simple expressions

$$
\begin{aligned}
\frac{\partial s_j}{\partial p_j} &= -\frac{\alpha}{1-\sigma_g} s_j \left(1-\sigma_g s_{j|g} - (1-\sigma_g) s_j\right) \\
\frac{\partial s_k}{\partial p_j} &= \frac{\alpha}{1-\sigma_g} s_j \left(\sigma_g s_{j|g} + (1-\sigma_g) s_k\right) \text{ if } k,j \text{ in same group} \\
&= \alpha s_j s_k \text{ if } k,j \text{ in different groups}
\end{aligned}
$$

- Simplifies to logit model if $\sigma_g = 0$

Random Coefficients (Mixed) Logit Model

- In order to relax logit and nested logit assumptions and constraints on substitution patterns and elasticities
- The linear Mixed Logit model is usually specified with

$$u_{ijt} = x_{jt}\beta_i + \alpha_i(I_i - p_{jt}) + \xi_{jt} + \varepsilon_{ijt}$$

where ε_{ijt} is a stochastic term with $Var(\varepsilon_{ijt}) = 1$

- 2 views of ε_{ijt}:
  - utility is deterministic, but the choice process itself is probabilistic (Tversky, 1972)
  - utility is deterministic, but ε_{ijt} captures the researcher's inability to formulate individual behavior precisely
- Interplay between ξ_{jt} and ε_{ijt}: in a way all the ξ_{jt} is doing is changing the mean of ε_{ijt}, by j and t.

Heterogeneity

- Consumer-level taste parameters are modeled as

$$\begin{aligned} \alpha_i &= \alpha + \sum_{r=1}^{d} \pi_{1r} D_{ir} + \sigma_1 v_{i1}, \\ \beta_{ik} &= \beta_k + \sum_{r=1}^{d} \pi_{(k+1)r} D_{ir} + \sigma_{k+1} v_{i(k+1)} \quad \text{for } k = 1, ..., K \end{aligned}$$

where

$D_i = (D_{i1}, ..., D_{id})'$ is a $d \times 1$ vector of observed demographic variables
$v_i = (v_{i1}, ..., v_{i(K+1)})'$ is a vector of $K+1$ unobserved consumer attributes
Denote Π is a $(K+1) \times d$ matrix of parameters and $\sigma = (\sigma_1, ..., \sigma_{K+1})$ is a vector of parameters.

- McFadden and Train (2000) shows that any random utility model can be approximated by a mixed logit model, provided the mixing distribution is adequate

Outside Option

The indirect utility from this outside option is

$$u_{i0t} = \alpha_i I_i + \varepsilon_{i0t}$$

- We normalize its mean utility to zero
- The outside option will allow for substitution outside the market; important in many IO applications
- Market level variation in the outside option may be important (trend in market size, ..)

Random Coefficient Logit

- Let $\theta = (\alpha, \beta, \Pi, \sigma)$ denote the parameters of the model.
  - $\theta_1 = (\alpha, \beta)$ the "linear" and $\theta_2 = (\Pi, \sigma)$ the "non-linear" parameters
- It will be convenient to write

$$u_{ijt} = \delta_{jt}(x_t, p_t, \xi_t; \alpha, \beta) + \mu_{ijt}(x_t, p_t, D_i; \Pi, \sigma) + \varepsilon_{ijt}$$

where $\delta_{jt} = x_{jt}\beta - \alpha p_{jt} + \xi_{jt}$ and

$$\begin{aligned} \mu_{ijt} &= -\left(\sum_{r=1}^{d} \pi_{1r} D_{ir} + \sigma_1 v_{i1}\right) p_{jt} + \\ &\quad \sum_{k=1}^{K} \left(\sum_{r=1}^{d} \pi_{(k+1)r} D_{ir} + \sigma_{k+1} v_{i(k+1)}\right) x_{jt}^k \end{aligned}$$

Random Coefficient Logit

$$u_{ijt} \equiv \delta(x_{jt}, p_{jt}, \xi_{jt}; \theta_1) + \mu(x_{jt}, p_{jt}, D_i, \nu_i; \theta_2) + \varepsilon_{ijt}$$

Note:

(1) mean utility will play a key role

(2) flexibility from interplay between μ_{ijt} and ε_{ijt}

(3) the "linear" and "non-linear" parameters

(4) definition of a market

Choice Probabilities and Market Shares

- Assume consumers purchase one unit of the good, which gives the highest utility.
- The probability that type (D_i, v_i) chooses option j is

$$\begin{aligned} s_{ijt} &= s_{ijt}(x_t, \delta_t, p_t, D_i, v_i; \theta) \\ &= P[u_{ijt} \geq u_{ikt} \ \forall k \mid x_t, \delta_t, p_t, D_i, v_i; \theta] \\ &= \int 1[u_{ijt} \geq u_{ikt} \ \forall k \mid x_t, \delta_t, p_t, D_i, v_i; \theta] dF_\varepsilon(\varepsilon). \end{aligned}$$

- We get aggregate market shares by integrating this probability over consumer attributes

$$s_{jt} = s_{jt}(x_t, \delta_t, p_t; \theta) = \int s_{ijt}(x_t, \delta_t, p_t, D_i, v_i; \theta) dF_D(D) dF_v(v)$$

- With assumptions on the distribution of the individual attributes F_v, one can compute this integral

Individual Choice Probabilities

- With consumer level data we will integrate only over v_i to obtain consumer level choice probabilities

$$s_{ijt} = s_{jt}(x_t, \delta_t, p_t, D_i; \theta) = \int s_{ijt}(x_t, \delta_t, p_t, D_i, v_i; \theta) dF_v(v|D_i)$$

where the distribution of unobserved v_i can depend on observed demographics.

- With assumptions on the distribution of the individual attributes $F_{v|D}$, one can compute this integral

Remarks: Logit

- Simple assumptions to recover the Logit model
  1. $\Pi = 0$ and $\sigma = 0$, which implies $\beta_i = \beta$ and $\alpha_i = \alpha$
  2. ε_{ijt} are i.i.d.
  3. ε_{ijt} are distributed according to a Type I extreme value distribution.
- These imply

$$s_{jt} = \frac{\exp\{x_{jt}\beta - \alpha p_{jt} + \xi_{jt}\}}{1 + \sum_{k=1}^{J} \exp\{x_{kt}\beta - \alpha p_{kt} + \xi_{kt}\}}$$

Price Elasticities: Logit

$$\eta_{jkt} = \frac{\partial s_{jt}}{\partial p_{kt}} \frac{p_{kt}}{s_{jt}} = \begin{cases} -\alpha p_{jt}(1 - s_{jt}) & \text{if } j = k \\ \alpha p_{kt} s_{kt} & \text{otherwise} \end{cases}$$

- 2 Problems
- Own price elasticities: market shares are small, so $\alpha(1 - s_{jt})$ is nearly constant and therefore the own-price elasticities are proportional to price.
  - driven mostly by lack of heterogeneity
- Cross-price elasticities: cross price elasticity w.r.t. a change in the price of product k is that same for all products such that $j \neq k$.
  - driven by lack of heterogeneity and i.i.d.

Remarks: Nested Logit

- Simple assumptions to recover the Nested Logit
  - $\Pi = 0$ and $\sigma = 0$,
  - divide the products into mutually exclusive nests, $g = 1, .., G$.
  - let $\varepsilon_{ijt} = \lambda \varepsilon_{ig(j)t} + \varepsilon^1_{ijt}$, where ε^1_{ijt} is an i.i.d. extreme value shock, $\varepsilon_{ig(j)t}$ is a shock common to all options in segment g, and λ is a parameter that captures the relative importance of the two.
  - a particular distribution for $\varepsilon_{ig(j)t}$ gives the Nested Logit model.
- The Nested Logit model is a private case of the more general Generalized Extreme Value model which imposes correlation among the options through correlation in ε_{ijt}.

Price Elasticities: Nested Logit

$$
\begin{aligned}
\eta_{jkt} &\equiv \frac{\partial s_{jt}}{\partial p_{kt}} \frac{p_{kt}}{s_{jt}} \\
&= -\frac{\alpha}{1-\sigma_g}\left(1-\sigma_g s_{jt|g}-(1-\sigma_g)s_{jt}\right) p_{jt} \qquad \text{if } j=k \\
&= \frac{\alpha}{1-\sigma_g}\left(\sigma_g s_{jt|g}+(1-\sigma_g)s_{kt}\right) p_{kt} \qquad \text{if } k\neq j \text{ but in same group} \\
&= \alpha s_{kt} p_{kt} \qquad \text{if } k\neq j \text{ and not in same group}
\end{aligned}
$$

Heterogeneity

- Correlations of μ_{ijt} allowing heterogeneity in tastes for the product attributes.
- For example, if "luxury" is an attribute of a car, then a consumer who likes one luxury car is more likely than the average consumer to like another luxury car
- Nested Logit can be seen as a private case

Price Elasticities: Random Coefficient Logit

$$\eta_{jkt} = \frac{\partial s_{jt}}{\partial p_{kt}} \frac{p_{kt}}{s_{jt}} = \begin{cases} -\frac{p_{jt}}{s_{jt}} \int \alpha_i s_{ijt}(1 - s_{ijt}) dP_D(D) dP_v(v) & \text{if } j = k \\ \frac{p_{kt}}{s_{jt}} \int \alpha_i s_{ijt} s_{ikt} dP_D(D) dP_v(v) & \text{otherwise} \end{cases}$$

Data Structures

- Market-level data
  - cross section/time series/panel of markets
- Consumer level data
  - cross section of consumers
  - sometimes: panel (i.e., repeated choices)
- Combination
  - sample of consumers plus market-level data
  - quantity/share by demographic groups
  - average demographics of purchasers of good j

Market-level Data

- We observe product-level quantity/market shares by "market"
- Data include:
  - aggregate (market-level) quantity
  - prices/characteristics/advertising
  - definition of market size
  - distribution of demographics
    - sample of actual consumers
    - data to estimate a parametric distribution
- Advantages:
  - easier to get
  - sample selection less of an issue
- Disadvantages
  - estimation often harder and identification less clear
  - need to specify the market size

Consumer-level Data

- Observe match between consumers and their choices
- Data include:
  - consumer choices (including choice of outside good)
  - consumer demographics
  - (eventually) prices/characteristics/advertising of *all* options
- Advantages:
  - impact of demographics
  - identification and estimation
  - dynamics (especially if we have panel)
- Disadvantages
  - harder/more costly to get
  - sample selection and reporting error

Random Coefficient (Mixed) Logit Model: Estimation

Specification

$$u_{ijt} = x_{jt}\beta_i + \alpha_i p_{jt} + \xi_{jt} + \varepsilon_{ijt}$$

where ε_{ijt} is i.i.d. extreme value distributed and

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

$$u_{ijt} \equiv \delta(x_{jt}, p_{jt}, \boldsymbol{\xi}_{jt}; \theta_1) + \mu(x_{jt}, p_{jt}, D_i, \nu_i; \theta_2) + \varepsilon_{ijt}$$

where $\delta_{jt} = x_{jt}\beta + \alpha p_{jt} + \xi_{jt}$, and $\mu_{ijt} = (p_{jt}\ x_{jt})\,(\Pi D_i + \Sigma v_i)$

Estimation with Micro Data

- Estimate in two steps.
- First step, estimate (δ, θ_2) say by MLE

  $$\Pr(y_{it} = j | D_{it}, \mathbf{x}_t, \mathbf{p}_t, \boldsymbol{\xi}_t, \theta) = \Pr(y_{it} = j | D_{it}, \delta(\mathbf{x}_t, \mathbf{p}_t, \boldsymbol{\xi}_t, \theta_1), \mathbf{x}_t, \mathbf{p}_t, \theta_2)$$

  For example assume ε_{ijt} is *iid* double exponential (Logit), and $\Sigma = 0$

  $$= \frac{\exp\{\delta_{jt} + (p_{jt} \;\; x_{jt})\Pi D_i\}}{\sum_{k=0}^{J} \exp\{\delta_{kt} + (p_{kt} \;\; x_{kt})\Pi D_i\}}$$

- Second step, recover θ_1

  $$\widehat{\delta}_{jt} = x_{jt}\beta + \alpha p_{jt} + \xi_{jt}$$

  ξ_{jt} is the residual. If it is correlated with price (or x's) need IVs (or an assumption about the panel structure)
- If $\Sigma \neq 0$, then we need to specify the distribution of ν_i

Intuition from estimation with consumer data

- Estimation in 2 steps: first recover δ (mean utilities) and θ_2 (parameters of heterogeneity) and then recover θ_1
- Different variations identifying the different parameters
  - θ_2 is identified from variation in demographics (D_i) holding the level (i.e., δ) constant
    - If $\Sigma \neq 0$ then it is identified from within market variation in choice probabilities
  - θ_1 is identified from cross market variation (and appropriate exclusion restrictions)
- With market-level data, try to follow a similar logic
  - however, we do not have within market variation to identify θ_2
  - will rely on cross market variation (in choice sets and demographics) for both steps

Mixed Logit Estimation with Simulated Maximum Likelihood

- Estimation by Simulated Maximum Likelihood with N observations (decision makers) is possible if sample is "exogenously" drawn.
- With

$$u_{ijt} = x_{jt}\beta_i + \alpha_i p_{jt} + \xi_{jt} + \varepsilon_{ijt}$$

$$\Pr(y_{it} = j|\theta_i) = \frac{\exp\{x_{jt}\beta_i + \alpha_i p_{jt} + \xi_{jt}\}}{\sum_{k=0}^{J} \exp\{x_{kt}\beta_i + \alpha_i p_{kt} + \xi_{kt}\}}$$

The conditional probability to observe a sequence of T choices is

$$S_i(\theta_i) = \prod_{t=1}^{T} \prod_{j} \Pr(y_{it} = j|\theta_i)^{d_{ijt}}$$

where $d_{ijt} = 1$ if i chooses j and 0 otherwise. $(\theta_i = (\alpha_i, \beta_i))$

Mixed Logit Estimation with Simulated Maximum Likelihood

- Denoting $f(\theta|\bar{\theta})$ the density of parameters θ, the unconditional probability is

$$\int \prod_{t=1}^{T} \prod_{j} \Pr(y_{it} = j|\theta_i)^{d_{ijt}} f(\theta_i|\bar{\theta}) d\bar{\theta}$$

- Then simulated log likelihood is

$$SLL(\bar{\theta}) = \sum_{i=1}^{N} \ln \left[\frac{1}{R} \sum_{r=1}^{R} S_i(\theta^r) \right]$$

where R is the number of replications and θ^r is the r^{th} draw from $f(\theta|\bar{\theta})$.

Key Challenges for Estimation with market level data

- Recovering the non-linear parameters θ_2, which govern heterogeneity, without observing consumer level data
- The unobserved characteristic, ξ_{jt}
  - generates a potential for correlation with price (or other x's)
  - when constructing a counterfactual we will have to deal with what happens to ξ_{jt}
- Consumer-level vs. Market-level data
  - with consumer data, the first issue is less of a problem
  - the "endogeneity" problem can exist with both consumer and market level data: a point often missed

Estimation with market level data

In principle, we could consider estimating θ by minimizing the distance between observed and predicted shares:

$$\min_{\theta} \left\| S_t - s_j(\mathbf{x}_t, \mathbf{p}_t, \theta) \right\|$$

- Issues:
  - computation (all parameters enter non-linearly)
  - prices might be correlated with the ξ_{jt} ("structural" error)

Inversion

- Instead, follow method proposed by Berry (1994) and BLP (1995)
- Key insight:
  - with ξ_{jt}, predicted shares can equal observed shares

$$\sigma_j(\delta_t, \mathbf{x}_t, \mathbf{p}_t; \theta_2) \equiv \int \mathbf{1}\left[u_{ijt} \geq u_{ikt} \quad \forall k \neq j\right] \, dF\left(\varepsilon_{it}, D_{it}, \nu_{it}\right) = S_{jt}$$

  - under weak conditions this mapping can be inverted

$$\delta_t = \sigma^{-1}(\mathbf{S}_t, \mathbf{x}_t, \mathbf{p}_t; \theta_2)$$

  - the mean utility is linear in ξ_{jt}; thus, we can form linear moment conditions
  - estimate parameters via GMM, using orthogonality between ξ_{jt} and instruments Z_{jt}

$$E\left(\xi_{jt} | Z_{jt}\right) = 0$$

Important point

- IVs play dual role
  - generate moment conditions to identify non linear parameters θ_2
  - deal with the correlation of prices and error
- Why different than consumer-level data?
  - with aggregate data we only know the mean choice probability, i.e., the market share
  - with consumer level data we know more moments of the distribution of choice probabilities (holding ξ_{jt} constant) : these moments help identify the heterogeneity parameters

Estimation with market data

- More frequent availability of market level data in IO
- Assume that we have valid IVs Z
  - later we will discuss where these come from
- Description of BLP algorithm

The BLP Estimation Algorithm

1. Compute predicted shares: given δ_t and θ_2 compute $\sigma_j(\delta_t, \mathbf{x}_t, \mathbf{p}_t; \theta_2)$
2. Inversion: given θ_2 search for δ_t that equates $\sigma_j(\delta_t, \mathbf{x}_t, \mathbf{p}_t; \theta_2)$ and the observed market shares S_{jt}
   - the search for δ_t will call the function computed in (1)
3. Use the computed δ_t to compute ξ_{jt} and form the GMM objective function (as a function of θ)
4. Search for the value of θ that minimizes the objective function

Example: Estimation of the Logit Model

- Data: aggregate quantity, price, characteristics. Market share $s_{jt} = q_{jt} / M_t$
  - Note: need for data on market size
- Computing market share

$$s_{jt} = \frac{\exp\{\delta_{jt}\}}{\sum_{k=0}^{J} \exp\{\delta_{kt}\}}$$

- Inversion

$$\ln(s_{jt}) - \ln(s_{0t}) = \delta_{jt} - \delta_{0t} = x_{jt}\beta + \alpha p_{jt} + \xi_{jt}$$

- Estimate using linear methods (e.g., 2SLS) with $\ln(s_{jt}) - \ln(s_{0t})$ as the "dependent variable".

Step 1: Compute the market shares predicted by the model

- Given δ_t and θ_2 (and the data) compute

$$\sigma_j(\delta_t, \mathbf{x}_t, \mathbf{p}_t; \theta_2) = \int \mathbf{1}\left[u_{ijt} \geq u_{ikt} \quad \forall k \neq j\right] dF\left(\varepsilon_{it}, D_{it}, \nu_{it}\right)$$

- For some models this can be done analytically (Logit, Nested Logit ..)
- Generally the integral is computed numerically
- A common way to do this is via simulation

$$\widetilde{\sigma}_j(\delta_t, \mathbf{x}_t, \mathbf{p}_t, F_{ns}; \theta_2) = \frac{1}{ns} \sum_{i=1}^{ns} \frac{\exp\{\delta_{jt} + (p_{jt}\ x_{jt})(\Pi D_i + \Sigma v_i)\}}{1 + \sum_{k=1}^{J} \exp\{\delta_{kt} + (p_{kt}\ x_{kt})(\Pi D_i + \Sigma v_i)\}}$$

where v_i and D_i, $i = 1, .., ns$ are draws from $F_v^*(v)$ and $F_D^*(D)$.

Step 2: Invert the shares to get mean utilities

- Given θ_2, for each market compute mean utility, δ_t, that equates the market shares computed in Step 1 to observed shares by solving

$$\widetilde{\sigma}(\delta_t, \mathbf{x}_t, \mathbf{p}_t, F_{ns}; \theta_2) = S_t$$

- For some model (e.g., Logit and Nested Logit) this inversion can be computed analytically.
- Generally solved using a contraction mapping for each market

$$\delta_t^{h+1} = \delta_t^h + \ln(S_t) - \ln(\widetilde{\sigma}(\delta_t^h, \mathbf{x}_t, \mathbf{p}_t, F_{ns}; \theta_2) \quad h = 0, ..., H,$$

where H is the smallest integer such that $\left\| \delta_t^H - \delta_t^{H-1} \right\| < \rho$

- δ_t^H is the approximation to δ_t
- Choosing a low tolerance level, ρ, is crucial (at least 10^{-12})

Step 3: Compute the GMM objective

- Once the inversion has been computed the error term is defined as

$$\xi_{jt}(\theta) = \widetilde{\sigma}^{-1}(\mathbf{S}_t, \mathbf{x}_t, \mathbf{p}_t; \theta_2) - x_{jt}\beta - \alpha p_{jt}$$

- Note: θ_1 enters linearly this term , and the GMM objective, while θ_2 enters non-linearly.
- This error is interacted with the IV to form

$$\xi(\theta)' ZWZ' \xi(\theta)$$

where W is the GMM weight matrix

Step 4: Search for the parameters that minimize the objective

- In general, the search is non-linear
- It can be simplified in two ways.
  - "concentrate out" the linear parameters and limit search to θ_2
  - use the Implicit Function Theorem to compute the analytic gradient and use it to aid the search
- Still highly non-linear so much care should be taken:
  - start search from different starting points
  - use different optimizers

Identification

- Ideal experiment: randomly vary prices, characteristics and availability of products, and see where consumers switch (i.e., shares of which products respond)
- In practice we will use IVs that try to mimic this ideal experiment
- Is there "enough" variation to identify substitution?
- Solutions:
  - supply information (BLP 1995)
  - many markets (Nevo 2001)
  - add micro information (MicroBLP 1994, Petrin 2003)

Instrumental Variables

- Discuss the estimation using market level data based on the moment condition
$$E(\xi_{jt}|z_{jt}) = 0.$$
using orthogonality between demand shock and IVs.
- IVs play dual role
  - generate moment conditions to identify θ_2
  - deal with the correlation of prices and error

Commonly used IVs: competition in characteristics space

- Assume that $E(\xi_{jt}|\mathbf{x}_t) = 0$, observed characteristics are mean independent of unobserved characteristics
- BLP propose using
  - own characteristics
  - sum of char of other products produced by the firm
  - sum of char of competitors products
- Power: proximity in characteristics space to other products $\longrightarrow$ markup $\longrightarrow$ price
- Validity: x_{jt} are assumed set before ξ_{jt} is known
- Most commonly used
  - do not require data we do not already have
- Often called "BLP Instruments"
- Not hard to come up with stories that make these invalid

Commonly used IVs: cost based

- Cost data are rarely directly observed
- BLP (1995, 1999) use characteristics that enter cost (but not demand)
- Villas-Boas (2007) or Bonnet and Dubois (2010) use prices of inputs interacted with product dummy variables (to generate variation by product)
- Hausman (1996) and Nevo (2001) rely on indirect measures of cost
    - use prices of the product in other markets
    - validity: after controlling for common effects, the unobserved characteristics are assumed independent across markets
    - power: prices will be correlated across markets due to common marginal cost shocks
    - easy to come up with examples where these IVs are not valid (national promotions, ..)

Commonly used IVs: dynamic panel

- Ideas from the dynamic panel data literature (Arellano and Bond, 1991, Blundell and Bond, 1998) have been used to motivate the use of lagged characteristics as instruments.
- Proposed in a footnote in BLP
- For example, Sweeting (2012) assumes $\xi_{jt} = \rho\xi_{jt-1} + \eta_{jt}$, where $E(\eta_{jt}|\mathbf{x}_{t-1}) = 0$.Then

$$E(\xi_{jt} - \rho\xi_{jt-1}|\mathbf{x}_{t-1}) = 0$$

is a valid moment condition

"Optimal " IVs

- Optimal instrumental variables (Chamberlain, 1987)

$$E\left[\frac{\partial \zeta_{jt}(\theta)}{\partial \theta'} \middle| \mathbf{x}_t\right]$$

- Use derivatives of mean utility with respect to coefficients and in particular to variance coefficients
- Approximate by taking derivatives at mean instead of mean of derivatives) (Berry, Levinsohn and Pakes, 1999, Reynaert and Verboven, 2014)

BLP Estimation with Matlab

Estimation with aggregate data

- Code using NFP algorithm provided by Nevo (2000), Knittel and Metaxoglou (2014)
- Available at http://web.mit.edu/knittel/www/KM_website.html
- Cereals data of Nevo (2000)
- 24 products (brands), 47 cities*2 quarters:94 markets, 24*94=2256 observations
- Variables: brand id, price, sugar content, mushy dummy
- Specification of demand model: cereal brand dummies which subsume product characteristics other than prices, unobserved consumer taste (random coefficient) interacted with constant, price, sugar content, mushy dummy

BLP Estimation with Matlab

Estimation with aggregate data

- Run main.m
- Run optim_results_summary.m
- Results in folder \Optimization results\

Mixed Multinomial Logit with Stata - Example 1

Estimation with micro level data

- Estimation with micro level data
- Data: teachers' evaluations of pupil behavior at school
  - one cross section: information about 1,313 pupils in 48 schools
  - Variables id and scy3 identify pupils and schools
  - Teachers group pupils in three different quality levels (tby), our dependent variable
  - Additional variables explaining the quality level of the pupils, such as sex
- See codes on Kenneth Train web page http://elsa.berkeley.edu/ train/software.html
- Stata code from: http://www.stata-journal.com/sjpdf.html?articlenum=st0133
- See Stata code mixlogit2.do

Mixed Multinomial Logit in Stata - Example 2

Estimation with micro level data

- Estimation with micro level data
- Data extract of Revelt and Train (1998) on households' choice of electricity supplier.
- Sample of residential electricity customers were presented four alternative electricity suppliers in 12 stated preference choice situations
- Characteristics: price per kilowatt-hour, length of contract, whether the company is local, and whether it is well known.
- Depending on experiment, price is either fixed or variable rate that depends on time of day or season.

Mixed Multinomial Logit Estimation in Stata - Example 2

Estimation with micro level data

- Model with following explanatory variables :
  - Price in cents per kilowatt-hour if fixed price, 0 if time-of-day or seasonal rates
  - Contract length in years
  - Whether company is local (0/1 dummy)
  - Whether company is well known (0/1 dummy)
  - Time-of-day rates (0/1 dummy)
  - Seasonal rates (0/1 dummy)
- See Stata code mixlogit3.do
- Data setup: Each observation corresponds to an alternative, and the dependent variable y is 1 for the chosen alternative in each choice situation and 0 otherwise.
- *gid* identifies the alternatives in a choice situation, *pid* identifies the decision maker, remaining variables are alternative attributes

Mixed Multinomial Logit Estimation in Stata - Example 2

Estimation with micro level data

- When modelling the random coefficient with log normal distribution $LN(\mu, \sigma^2)$ (means the log of coefficient is $N(\mu, \sigma^2)$), then the median, mean and standard deviation of coefficient of price with log-normal distribution $LN(\mu, \sigma^2)$ are:

$$\begin{aligned} \text{median} \quad & \exp \mu \\ \text{mean} \quad & \exp(\mu + \frac{\sigma^2}{2}) \\ \text{standard deviation} \quad & \exp(\mu + \frac{\sigma^2}{2})\sqrt{\exp(\sigma^2) - 1} \end{aligned}$$

Conclusion

- Extensions of demand models:
  - Dynamics, repeated choices: state dependence vs. unobserved heterogeneity
  - Robustness to unobserved characteristic bias using control functions (Petrin and Train, 2010)
  - Multiple Discreteness: Hendel (RES, 99) models the demand for computers by firms. Each firm buys several brands and several units.
  - Discrete/continuous models (Dubin-McFadden, 1984, Hanemann, 1984, Smith, 2004, Dubois and Jodar, 2011)

BLP, 1995

- Berry, Levinsohn, Pakes "Automobile Prices in Market Equilibrium" (1995)

1. The effect of IV
2. Logit versus RC Logit

Data

- 20 years of annual US national data, 1971-90 (T=20): 2217 model-years
- Quantity data by name plate (excluding fleet sales)
- Prices – list prices
- Characteristics from Automotive News Market Data Book
- Price and characteristics correspond to the base model
- Market Size: number of households in the US

Demand Model

The indirect utility is

$$u_{ijt} = x_{jt}\beta_i + \alpha \ln(y_i - p_{jt}) + \xi_{jt} + \varepsilon_{ijt}$$

Note: income enters differently than before
Use random coefficient β_i and distribution of y_i in the CPS

$$\beta_i^k = \beta^k + \sigma^k v_{ik} \quad v_{ik} \sim N(0,1)$$

The outside option has utility

$$u_{i0t} = \alpha \ln(y_i) + \xi_{0t} + \sigma^0 v_{i0} + \varepsilon_{i0t}$$

Estimation

- Basically estimate as we discussed before:
  - add supply-side moments (changes last step of the algorithm)
    - help pin down demand parameters
    - adds cost side IVs
  - Instrumental variables. assume $E(\xi_{jt}|\mathbf{x}_t) = 0$, and use
    - (i) own characteristics
    - (ii) sum of char of other products produced by the firm
    - (iii) sum of characteristics products produced by other firms
  - Cost side IVs: $E(\xi_{jt}|\mathbf{w}_t) = 0$

Table 3: Effect of IV (in Logit)

TABLE III

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

Notes: The standard errors are reported in parentheses.
* The continuous product characteristics—hp/weight, size, and fuel efficiency (*MP$* or *MPG*)—enter the demand equations in levels, but enter the column 3 price regression in natural logs.

Table 4: Demand Estimates

TABLE IV

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

Interpretation

- Positive coefficient on HP/Weight:
  - Either high mean for distribution of tastes
  - Or large variance of that distribution, or both
- Explanations have different implications for substitution patterns, and how market share will change with product attributes and prices
- If zero standard deviation: when a high HP/weight car increases its price, consumers who substitute away from that car have the same marginal utilities for HP/weight as any other consumer and will not tend to substitute disproportionately toward other high HP/weight cars.
- If, standard deviation of tastes for HP/weight relatively large, consumers who substitute away will be consumers who placed a high marginal utility on HP/weight originally, and should substitute disproportionately towards other high HP/weight cars.

Tables 5: Elasticities

TABLE V

A SAMPLE FROM 1990 OF ESTIMATED DEMAND ELASTICITIES WITH RESPECT TO ATTRIBUTES AND PRICE (BASED ON TABLE IV (CRTS) ESTIMATES)

Model	Value of Attribute/Price Elasticity of demand with respect to: HP/Weight	Air	MP$	Size	Price
Mazda323	0.366	0.000	3.645	1.075	5.049
	0.458	0.000	1.010	1.338	6.358
Sentra	0.391	0.000	3.645	1.092	5.661
	0.440	0.000	0.905	1.194	6.528
Escort	0.401	0.000	4.022	1.116	5.663
	0.449	0.000	1.132	1.176	6.031
Cavalier	0.385	0.000	3.142	1.179	5.797
	0.423	0.000	0.524	1.360	6.433
Accord	0.457	0.000	3.016	1.255	9.292
	0.282	0.000	0.126	0.873	4.798
Taurus	0.304	0.000	2.262	1.334	9.671
	0.180	0.000	−0.139	1.304	4.220
Century	0.387	1.000	2.890	1.312	10.138
	0.326	0.701	0.077	1.123	6.755
Maxima	0.518	1.000	2.513	1.300	13.695
	0.322	0.396	−0.136	0.932	4.845
Legend	0.510	1.000	2.388	1.292	18.944
	0.167	0.237	−0.070	0.596	4.134
TownCar	0.373	1.000	2.136	1.720	21.412
	0.089	0.211	−0.122	0.883	4.320
Seville	0.517	1.000	2.011	1.374	24.353
	0.092	0.116	−0.053	0.416	3.973
LS400	0.665	1.000	2.262	1.410	27.544
	0.073	0.037	−0.007	0.149	3.085
BMW 735i	0.542	1.000	1.885	1.403	37.490
	0.061	0.011	−0.016	0.174	3.515

Tables 6: Own and cross price semi-elasticities

TABLE VI

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.

Table 7: Substitution to the outside option

TABLE VII

SUBSTITUTION TO THE OUTSIDE GOOD

Model	Given a price increase, the percentage who substitute to the outside good (as a percentage of all who substitute away.) Logit	BLP
Mazda 323	90.870	27.123
Nissan Sentra	90.843	26.133
Ford Escort	90.592	27.996
Chevy Cavalier	90.585	26.389
Honda Accord	90.458	21.839
Ford Taurus	90.566	25.214
Buick Century	90.777	25.402
Nissan Maxima	90.790	21.738
Acura Legend	90.838	20.786
Lincoln Town Car	90.739	20.309
Cadillac Seville	90.860	16.734
Lexus LS400	90.851	10.090
BMW 735*i*	90.883	10.101

Markups

- The profits of firm f

$$\Pi_f = \sum_{j \in \mathcal{F}_f} (p_j - mc_j) M s_j(p)$$

the first order conditions are

$$s_j(p) + \sum_{r \in \mathcal{F}_f} (p_r - mc_r) \frac{\partial s_r(p)}{\partial p_j} = 0$$

Define $\Delta_{jr} = -\partial s_r / \partial p_j \quad j, r = 1, ..., J$, and
$\Omega_{jr} = \begin{cases} \Delta_{jr} & \text{if } \exists \{r, j\} \subset \mathcal{F}_f \\ 0 & \text{otherwise} \end{cases}$

$$s(p) + \Omega(p - mc) = 0$$

and

$$p - mc = \Omega^{-1} s(p)$$

Table 8: Markups

TABLE VIII

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

Summary

- Powerful method with potential for many applications
- Clearly show:
  - effect of IV
  - RC logit versus logit
- Common complaints:
  - instruments
  - supply side: static, not tested, driving the results
  - demand side dynamics

Nevo, "Measuring Market Power in the Ready-to-eat Cereal Industry" (2001)

1. Effects of various IV's
2. Testing the model of competition
3. Comparison to alternative demand models

The RTE cereal industry

- Characterized by:
  - high concentration ($C3 \approx 75\%$, $C6 \approx 90\%$) (m largest firms hold market share Cm)
  - high price-cost margins ($\approx 45\%$)
  - large advertising to sales ratios ($\approx 13\%$)
  - numerous introductions of brands (67 new brands by top 6 in 80's)
- This has been used to claim that this is a perfect example of collusive pricing

Questions

- Is pricing in the industry collusive?
- What portion of the markups in the industry due to:
    - Product differentiation?
    - Multi-product firms?
    - Potential price collusion?

Strategy

- Estimate brand level demand
- Compute PCM predicted by different industry structures\models of conduct:
  - Single-product firms
  - Current ownership (multi-product firms)
  - Fully collusive pricing (joint ownership)
- Compare predicted PCM to observed PCM

Supply

- The profits of firm f

$$\Pi_f = \sum_{j \in \mathcal{F}_f} (p_j - mc_j) q_j(p) - C_f$$

the first order conditions are *in expectations*

$$s_j(p) + \sum_{r \in \mathcal{F}_f} (p_r - mc_r) \frac{\partial s_r(p)}{\partial p_j} = 0$$

- Define $\Omega_{jr} = \begin{cases} \partial s_r / \partial p_j & \text{if } \{r, j\} \subset \mathcal{F}_f \\ 0 & \text{otherwise} \end{cases}$
- Then

$$s(p) + \Omega(p - mc) = 0$$

Supply

- Then

$$p - mc = -\Omega^{-1} s(p)$$

- Therefore by:
  1. assuming a model of conduct
  2. using estimates of demand substitution
- We can compute price-cost margins under different "ownership" structures

Example

- Firm A owns goods 1 and 2 and firm B owns good 3

$$\frac{\partial s}{\partial p} = \begin{bmatrix} \frac{\partial s_1}{\partial p_1} = -2.4 & \frac{\partial s_2}{\partial p_1} = 0.2 & \frac{\partial s_3}{\partial p_1} = 0.1 \\ \frac{\partial s_1}{\partial p_2} = 0.3 & \frac{\partial s_2}{\partial p_2} = -3.1 & \frac{\partial s_3}{\partial p_2} = 0.4 \\ \frac{\partial s_1}{\partial p_3} = 0.6 & \frac{\partial s_2}{\partial p_3} = 0.7 & \frac{\partial s_3}{\partial p_3} = -1.8 \end{bmatrix} \qquad s = \begin{bmatrix} 0.40 \\ 0.12 \\ 0.24 \end{bmatrix}$$

then

$$\begin{aligned} \begin{bmatrix} p_1 - mc_1 \\ p_2 - mc_2 \\ 0 \end{bmatrix} &= -\begin{bmatrix} -2.4 & 0.2 & 0 \\ 0.3 & -3.1 & 0 \\ 0 & 0 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 0.40 \\ 0.12 \\ 0.24 \end{bmatrix} = \begin{bmatrix} 0.17 \\ 0.05 \\ 0 \end{bmatrix} \\ p_3 - mc_3 &= 0.24/1.8 = 0.13 \end{aligned}$$

Demand

- Utility, as before

$$u_{ijt} = x_{jt}\beta_i + \alpha_i p_{jt} + \xi_{jt} + \varepsilon_{ijt}$$

- Allow for brand dummy variables (to capture the part of ξ_{jt} that does not vary by market)
  - captures characteristics that do not vary over markets
- Random coefficient logit on aggregate data (BLP) with only demand side moments

Data

- IRI Infoscan scanner data
  - market shares – defined by converting volume to servings
  - prices – pre-coupon real transaction per serving price
  - 25 brands (top 25 in last quarter), in 67 cities (number increases over time) over 20 quarters (1988-1992); 1124 markets, 27,862 observations
- LNA advertising data
- Characteristics from cereal boxes
- Demographics from March CPS
- Cost instruments from Monthly CPS
- Market size – one serving per consumer per day

Estimation

- Various specifications with OLS:
  - product characteristics (*i*)
  - brand dummy variables as controls (*ii*) and (*iii*) control for unobserved quality (instead of instrumenting for it)
- Explores various IVs:
  - brand dummy variables (*iv*)
  - average regional prices (*v*)
  - proxies for city level costs(density, earnings in retail sector, transportation costs) (*vi*)
  - addition of demographics (*vii*,*viii*,*ix*) increases absolute value of price coefficient
  - allows for city-specific intercepts to control city-level demand shocks (*x*)
- Identify taste coefficients by minimum distance in second stage

Logit Demand

TABLE V

RESULTS FROM LOGIT DEMAND[a]

Variable	OLS (i)	(ii)	(iii)	IV (iv)	(v)	(vi)	(vii)	(viii)	(ix)	(x)
Price	−4.96	−7.26	−7.97	−8.17	−17.57	−17.12	−22.56	−23.77	−23.37	−23.07
	(0.10)	(0.16)	(0.15)	(0.11)	(0.50)	(0.49)	(0.51)	(0.53)	(0.47)	(1.17)
Advertising	0.158	0.026	0.026	0.157	0.020	0.020	0.018	0.017	0.018	0.013
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002
Log of Median Income	—	—	0.89	—	—	—	1.06	1.13	1.12	—
			(0.02)				(0.02)	(0.02)	(0.02)	
Log of Median Age	—	—	−0.423	—	—	—	−0.063	0.003	−0.007	—
			(0.052)				(0.059)	(0.062)	(0.061)	
Median HH Size	—	—	−0.126	—	—	—	−0.053	−0.036	−0.038	—
			(0.027)				(0.029)	(0.031)	(0.031)	
Fit/Test of Over Identification[b]	0.54	0.72	0.74	436.9	168.5	181.2	83.96	82.95	85.87	15.06
				(26.30)	(30.14)	(16.92)	(30.14)	(16.92)	(42.56)	(42.56)
1st Stage R^2	—	—	—	0.889	0.908	0.908	0.910	0.909	0.913	0.952
1st Stage F-test	—	—	—	5119	124	288	129	291	144	180
Instruments[c]	—	—	—	brand dummies	prices	cost	prices	cost	prices, cost	prices, cost

Results from the Full Model

TABLE VI

RESULTS FROM THE FULL MODEL[a]

Variable	Means (β's)	Standard Deviations (σ's)	Interactions with Demographic Variables: Income	Income Sq	Age	Child
Price	−27.198	2.453	315.894	−18.200	—	7.634
	(5.248)	(2.978)	(110.385)	(5.914)		(2.238)
Advertising	0.020	—	—	—	—	—
	(0.005)					
Constant	−3.592[b]	0.330	5.482	—	0.204	—
	(0.138)	(0.609)	(1.504)		(0.341)	
Cal from Fat	1.146[b]	1.624	—	—	—	—
	(0.128)	(2.809)				
Sugar	5.742[b]	1.661	−24.931	—	5.105	—
	(0.581)	(5.866)	(9.167)		(3.418)	
Mushy	−0.565[b]	0.244	1.265	—	0.809	—
	(0.052)	(0.623)	(0.737)		(0.385)	
Fiber	1.627[b]	0.195	—	—	—	−0.110
	(0.263)	(3.541)				(0.0513)
All-family	0.781[b]	0.1330	—	—	—	
	(0.075)	(1.365)				
Kids	1.021[b]	2.031	—	—	—	
	(0.168)	(0.448)				
Adults	1.972[b]	0.247	—	—	—	
	(0.186)	(1.636)				
GMM Objective (degrees of freedom)			5.05 (8)			
MD χ^2			3472.3			
% of Price Coefficients > 0			0.7			

Elasticities

MEDIAN OWN AND CROSS-PRICE ELASTICITIES[a]

#	Brand	Corn Flakes	Frosted Flakes	Rice Krispies	Froot Loops	Cheerios	Total	Lucky Charms	P Raisin Bran	CapN Crunch	Shredded Wheat
1	K Corn Flakes	−3.379	0.212	0.197	0.014	0.202	0.097	0.012	0.013	0.038	0.028
2	K Raisin Bran	0.036	0.046	0.079	0.043	0.145	0.043	0.037	0.057	0.050	0.040
3	K Frosted Flakes	0.151	−3.137	0.105	0.069	0.129	0.079	0.061	0.013	0.138	0.023
4	K Rice Krispies	0.195	0.144	−3.231	0.031	0.241	0.087	0.026	0.031	0.055	0.046
5	K Frosted Mini Wheats	0.014	0.024	0.052	0.043	0.105	0.028	0.038	0.054	0.045	0.033
6	K Froot Loops	0.019	0.131	0.042	−2.340	0.072	0.025	0.107	0.027	0.149	0.020
7	K Special K	0.114	0.124	0.105	0.021	0.153	0.151	0.019	0.021	0.035	0.035
8	K Crispix	0.077	0.086	0.114	0.034	0.181	0.085	0.030	0.037	0.048	0.043
9	K Corn Pops	0.013	0.109	0.034	0.113	0.058	0.025	0.098	0.024	0.127	0.016
10	GM Cheerios	0.127	0.111	0.152	0.034	−3.663	0.085	0.030	0.037	0.056	0.050
11	GM Honey Nut Cheerios	0.033	0.192	0.058	0.123	0.094	0.034	0.107	0.026	0.162	0.024
12	GM Wheaties	0.242	0.169	0.175	0.025	0.240	0.113	0.021	0.026	0.050	0.043
13	GM Total	0.096	0.108	0.087	0.018	0.131	−2.889	0.017	0.017	0.029	0.029
14	GM Lucky Charms	0.019	0.131	0.041	0.124	0.073	0.026	−2.536	0.027	0.147	0.020
15	GM Trix	0.012	0.103	0.031	0.109	0.056	0.026	0.096	0.024	0.123	0.016
16	GM Raisin Nut	0.013	0.025	0.042	0.035	0.089	0.040	0.031	0.046	0.036	0.027
17	GM Cinnamon Toast Crunch	0.026	0.164	0.049	0.119	0.089	0.035	0.102	0.026	0.151	0.022
18	GM Kix	0.050	0.279	0.070	0.101	0.106	0.056	0.088	0.030	0.149	0.025
19	P Raisin Bran	0.027	0.037	0.068	0.044	0.127	0.035	0.038	−2.496	0.049	0.036
20	P Grape Nuts	0.037	0.049	0.088	0.042	0.165	0.050	0.037	0.051	0.052	0.047
21	P Honey Bunches of Oats	0.100	0.098	0.104	0.022	0.172	0.109	0.020	0.024	0.038	0.033
22	Q 100% Natural	0.013	0.021	0.046	0.042	0.103	0.029	0.036	0.052	0.046	0.029
23	Q Life	0.077	0.328	0.091	0.114	0.137	0.046	0.096	0.023	0.182	0.029
24	Q CapN Crunch	0.043	0.218	0.064	0.124	0.101	0.034	0.106	0.026	−2.277	0.024
25	N Shredded Wheat	0.076	0.082	0.124	0.037	0.210	0.076	0.034	0.044	0.054	−4.252
26	Outside good	0.141	0.078	0.084	0.022	0.104	0.041	0.018	0.021	0.033	0.021

Margins

- Although mean price sensitivity of full model similar to Logit model, implied markups are different (better estimation of cross-price elasticities in full model)
- Logit model: can use estimates to compute PCM for brands not included in estimation: just need price sensitivity and the market shares of additional brands. Not possible in full model.
- Test of model by comparing PCM with accounting margins: Bertrand Nash Pricing not rejected while collusion is rejected

Margins

TABLE III

DETAILED ESTIMATES OF PRODUCTION COSTS

Item	$/lb	% of Mfr Price	% of Retail Price
Manufacturer Price	2.40	100.0	80.0
Manufacturing Cost:	1.02	42.5	34.0
Grain	0.16	6.7	5.3
Other Ingredients	0.20	8.3	6.7
Packaging	0.28	11.7	9.3
Labor	0.15	6.3	5.0
Manufacturing Costs (net of capital costs)[a]	0.23	9.6	7.6
Gross Margin		57.5	46.0
Marketing Expenses:	0.90	37.5	30.0
Advertising	0.31	13.0	10.3
Consumer Promo (mfr coupons)	0.35	14.5	11.7
Trade Promo (retail in-store)	0.24	10.0	8.0
Operating Profits	0.48	20.0	16.0

[a] Capital costs were computed from ASM data.
Source: Cotterill (1996) reporting from estimates in CS First Boston Reports "Kellogg Company," New York, October 25, 1994.

Margins

TABLE VIII

MEDIAN MARGINS[a]

	Logit (Table V column ix)	Full Model (Table VI)
Single Product Firms	33.6% (31.8%–35.6%)	35.8% (24.4%–46.4%)
Current Ownership of 25 Brands	35.8% (33.9%–38.0%)	42.2% (29.1%–55.8%)
Joint Ownership of 25 Brands	41.9% (39.7%–44.4%)	72.6% (62.2%–97.2%)
Current Ownership of All Brands	37.2% (35.2%–39.4%)	—
Monopoly/Perfect Price Collusion	54.0% (51.1%–57.3%)	—

Comments

- Ignores the retailer – uses retailer prices to study manufacturer competition
    - retail margins go into marginal cost (unlike Villas Boas, 2007 or Bonnet and Dubois, 2010)
    - marginal costs do not vary with quantity, therefore this restricts the retailers pricing behavior
        - which direction will this bias the finding? Most likely towards finding collusion where there is none (the retailer behavior might take into account effects across products)
    - Villas Boas (2007), Bonnet and Dubois (2010) extend the model adding strategic retailer behavior

Comments

- Much of the price variation at the store-level is coming from "sales". How does this impact the estimation?
  - data is quite aggregated:quarter-brand-city
  - "sales" generate incentives for consumer to stockpile (Hendel and Nevo)
- Is choice discrete? Hanemann (1984), Smith (2004), Dubois and Jodar (2011) for discrete/continuous choice model.

Nevo, "Mergers with Differentiated Products: The Case of the Ready-to-Eat Cereal Industry" (2000)

- Demand estimation and use of a model of postmerger conduct to simulate competitive effects of a merger.
  - Estimate a brand-level demand system for ready-to-eat cereal with supermarket scanner data
  - Recover marginal costs
  - Simulate postmerger price equilibria
  - Compute welfare effects

Merger Simulation

- The profits of firm f

$$\Pi_f = \sum_{j \in \mathcal{F}_f} (p_j - mc_j) q_j(p) - C_f$$

the first order conditions are

$$s_j(p) + \sum_{r \in \mathcal{F}_f} (p_r - mc_r) \frac{\partial s_r(p)}{\partial p_j} = 0$$

- Defining $\Omega_{jr}^{pre} = \begin{cases} \frac{\partial s_r}{\partial p_j} & \text{if } \{r, j\} \subset \mathcal{F}_f \\ 0 & \text{otherwise} \end{cases}$ we can recover marginal costs

$$mc = p^{pre} + \Omega^{pre^{-1}} s(p^{pre})$$

Merger Simulation

- Then, post merger prices satisfy

$$p^{post} = mc - {\Omega^{post}}^{-1} s(p^{post})$$

that is

$$p^{post} = p^{pre} + {\Omega^{pre}}^{-1} s(p^{pre}) - {\Omega^{post}}^{-1} s(p^{post})$$

Merger Simulation

- Consumer welfare change can be computed with structural model, assuming no changes in the utility from the outside good and no change in unobserved components ξ_{jt}
- Consumer surplus for individual i is (Rosen (1981))

$$\begin{aligned} E\left[CS_i\right] &= \frac{1}{\alpha_i} E_\varepsilon\left[\max_j u_{ijt}\right] = \frac{1}{\alpha_i} E_\varepsilon\left[\max_j \left(x_{jt}\beta_i + \alpha_i p_{jt} + \xi_{jt} + \varepsilon_{ijt}\right)\right] \\ &= \frac{1}{\alpha_i} \ln \sum\nolimits_j \exp\left[x_{jt}\beta_i + \alpha_i p_{jt} + \xi_{jt}\right] + C \end{aligned}$$

- With constant marginal utility of income,

$$\Delta E\left[CS\right] = \int \frac{1}{\alpha_i} \left(\begin{array}{c} \ln \sum_j \exp\left[x_{jt}\beta_i + \alpha_i p_{jt}^{post} + \xi_{jt}\right] \\ - \ln \sum_j \exp\left[x_{jt}\beta_i + \alpha_i p_{jt}^{pre} + \xi_{jt}\right] \end{array} \right) dF\left(\alpha_i, \beta_i\right)$$

Merger Simulation

TABLE 7 **Change in Variable Profits and Consumer Surplus as a Result of Mergers (millions of dollars per year)**

	Post and Nabisco		General Mills and Nabisco	
Consumer surplus	−13.98		−26.79	
Profits/revenues (total)	6.20	−4.77	10.66	−12.33
Kellogg	2.56	3.77	5.54	7.57
General Mills	2.34	3.65	2.63	−7.50
Post	.60	−5.17	1.54	2.94
Quaker Oats	.54	.84	1.43	2.07
Ralston	.14	.25	.30	.52
Nabisco	.01	−8.11	−.77	−17.93
Total Welfare	−7.78		−16.13	

Merger Simulation

Cost reduction				
(so total welfare is unchanged)		1.5%		10.8%
Profits/revenues (total)	8.29	−1.81	16.89	−3.36
Kellogg	1.39	1.90	3.77	4.93
General Mills	1.35	1.92	.47	−13.46
Post	3.73	−.57	.65	1.18
Quaker Oats	.31	.43	1.12	1.58
Ralston	.09	.15	.20	.36
Nabisco	1.42	−5.65	10.68	2.07

The top half of the table is based on the results of Table 5. The bottom half displays the cost reductions required to keep total welfare unchanged, i.e., change in consumer surplus minus change in variable profits equals zero. The first three columns assume a fixed proportional reduction only to brands of acquired firm, while the last two columns assume cost reductions to brands of both firms.

Gowrisankaran, Nevo, Town (2015)

- Gowrisankaran, Nevo, Town (2015) "Mergers When Prices Are Negotiated: Evidence from the Hospital Industry." American Economic Review
- In many markets, prices are negotiated through bilateral negotiations.
- Bargaining ability and merging: A party in negotiations will earn more beneficial terms of trade by improving its bargaining position. One way a firm can increase its bargaining power is by merging with a competitor.

Merger when prices are negotiated

- Gowrisankaran, Nevo, Town (2015) estimate a model of competition in which prices are negotiated between Managed Care Organizations (MCOs) and Hospitals.
- Estimates are used to investigate the impact of counterfactual hospital mergers.
- The model is used to evaluate how hospital bargain and patient coinsurance restraint prices.

Model

The model of competition between MCOs and hospitals has two stages:

- First-stage: MCOs and hospital systems negotiate the base price that each hospital will be paid by each MCO for hospital care.
- Second-stage: Enrollees who are ill decide where to seek treatment, choosing a hospital to maximize utility.

Hospitals seek to maximize a weighted sum of profits and MCOs act as agents for self-insured employers, seeking to maximize a weighted sum of enrollee welfare and insurer costs.

Model

Second-stage: patient hospital choice

1. Patient i is stricken by illness $d = 0, 1, .., D$ with probability f_{id} ($d = 0$ no illness possibility).
2. For each realized illness d, the patient seeks hospital care at the hospital which gives him the highest utility (multinomial logit utility).

Model

Second-stage: patient hospital choice

The utility that patient i with illness d receives from care at hospital $j \in \mathcal{N}_{m(i)}$ is given by

$$u_{ijd} = \underbrace{\beta X_{ijd} - \alpha \overbrace{c_{id} w_d p_{m(i)j}}^{\text{out-of-pocket price}}}_{\delta_{ijd}} + \epsilon_{ijd}$$

where,

- $\mathcal{N}_{m(i)}$: set of hospitals in the network of MCO $m(i)$.
- c_{id}: individual i's coinsurance rate.
- w_d: relative intensity of resource use for illness d.
- $p_{m(i)j}$: negotiated price between i's MCO $m(i)$ and hospital j.

Outside option

$$u_{i0d} = -\alpha c_{id} w_d p_{m(i)0} + \epsilon_{i0d}$$

Model

Second-stage: patient hospital choice

- The choice probability for patient i with disease d as a function of prices and network structure is:

$$s_{ijd}(\mathcal{N}_{m(i)}, \mathbf{p}_{m(i)}) = \frac{\exp(\delta_{ijd})}{\sum_{k \in \mathcal{N}_{m(i)} \cup \{0\}} \exp(\delta_{ikd})}$$

- The ex-ante expected utility to patient i, as a function of illness probabilities, prices and the network of hospitals in the plan, is given by

$$W_i(\mathcal{N}_{m(i)}, \mathbf{p}_{m(i)}) = \sum_d f_{id} \ln \left(\sum_{k \in \mathcal{N}_{m(i)}} \exp(\delta_{ikd}) \right)$$

Model

First-stage: MCO and hospital bargaining

MCOs and hospital systems negotiate the base price that each hospital will be paid by each MCO for hospital care. There are $M \times S$ potential contracts, each specifying the negotiated base prices for one MCO m - hospital system s pair.

- Contract specifies all negotiated base prices for its pair (m, s).
- Each hospital j within a system s has a separate base price.

The outcome of these negotiations is determined via Horn and Wolinsky (1988 RAND) model:

- Given $\mathbf{p}_{m,-s}$, each contract solves price vector $\mathbf{p}_{m,s}$ as Nash bargaining solution.
- Disagreement points are second-stage values to each party from having no agreement.
- The full vector of prices corresponds to the Nash equilibrium of the simultaneous bilateral negotiations game.

Model

First-stage: MCO and hospital bargaining

The MCO pays the part of the bill that is not paid by the patient:

$$TC_m(\mathcal{N}_m, \mathbf{p}_m) = \sum_i \sum_d 1_{\{m(i)=m\}}(1 - c_{id}) f_{id} w_d \sum_{j \in \mathcal{N}_m \cup \{0\}} p_{mj} s_{ijd}(\mathcal{N}_m, \mathbf{p}_m)$$

The value for the MCO and the employer it represents is

$$V_m(\mathcal{N}_m, \mathbf{p}_m) = \frac{\tau}{\alpha} \sum\nolimits_i 1_{\{m(i)=m\}} W_i(\mathcal{N}_m, \mathbf{p}_m) - TC_m(\mathcal{N}_m, \mathbf{p}_m)$$

where τ is the relative weight on employee welfare. $\tau > 1$ means MCOs care more about enrollee welfare than insurer costs.
For any system s (with a set J_s of hospitals), the net value that MCO m receives from including system s in its network is:

$$V_m(\mathcal{N}_m, \mathbf{p}_m) - V_m(\mathcal{N}_m \setminus \mathcal{J}_s, \mathbf{p}_m)$$

Model

First-stage: MCO and hospital bargaining

The marginal cost of hospital j for treating an illness with weight $w_d = 1$ from a patient at MCO m is:

$$mc_{mj} = \gamma v_{mj} + \varepsilon_{mj}$$

The normalized quantity to hospital $j \in \mathcal{N}_m$ is:

$$q_{mj}(\mathcal{N}_m, \mathbf{p}_m) = \sum_i \sum_d 1_{\{m(i)=m\}} f_{id} w_d s_{ijd}(\mathcal{N}_m, \mathbf{p}_m)$$

The net value that system s receives from including MCO m is:

$$\Pi_{m,s}(\mathcal{N}_m, \mathbf{p}_m) = \sum_{j \in \mathcal{J}_s} q_{mj}(\mathcal{N}_m, \mathbf{p}_m)[p_{mj} - mc_{mj}]$$

The Nash bargaining problem for MCO m and system s is the Nash product of the net values from agreement:

$$NB_{m,s}(\mathbf{p}_{m,s}|\mathbf{p}_{m,-s}) = \Pi_{m,s}(\mathcal{N}_m, \mathbf{p}_m)^{b_{s(m)}} [V_m(\mathcal{N}_m, \mathbf{p}_m)\text{-}V_m(\mathcal{N}_m \backslash \mathcal{J}_s, \mathbf{p}_m)]^{b_{m(s)}}$$

Model

First-stage: MCO and hospital bargaining

Let $\mathbf{p}^*_{m,s}$ denote the Horn and Wolinsky (1988 RAND) price vector for MCO m for all hospitals $j \in \mathcal{J}_s$. It must satisfy the Nash bargain for each contract, conditioning on the outcomes of other contracts. Thus, each p^*_{mj} satisfies

$$p^*_{mj} = \arg\max_{p_{mj}} NB_{m,s}(p_{mj}, \mathbf{p}_{m,-j} \mid \mathbf{p}^*_{m,-s})$$

Solving for the equilibrium prices yields:

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

where

$$\Omega(j,k) = \frac{\partial q_{mk}}{\partial p_{mj}} \quad \text{and} \quad \Lambda(j,k) = \frac{b_{m(s)}}{b_{s(m)}} \frac{\frac{\partial V_m(\mathcal{N}_m, \mathbf{p}_m)}{\partial p_{mj}} q_{mk}}{V_m(\mathcal{N}_m, \mathbf{p}_m) - V_m(\mathcal{N}_m \backslash \mathcal{J}_s, \mathbf{p}_m)}$$

Model

First-stage: MCO and hospital bargaining

- The previous formula shows we can back out implied marginal costs for the bargaining model as a closed-form function of prices, quantities and derivatives, given MCO and patient incentives.
- Formula is analogous to the standard Lerner index equation, but where actual patient price sensitivity is replaced by the effective price sensitivity of the MCO. If $\Lambda = 0$, this would be Bertrand competition. $\Omega + \Lambda$ is the price sensitivity.
    - With identical hospital, bargaining lower equilibrium prices compared to Bertrand
    - But with asymmetric hospitals and multi-hospital systems, there is a steering effect which may not uniformly lower prices.

Model

- Impact of price on MCO surplus $V_m(\mathcal{N}_m, \mathbf{p}_m)$ is ambiguous
- With $\tau = 1$:

$$\frac{\partial V_m}{\partial p_{mj}} = -q_{mj}$$
$$-\alpha \sum_i \sum_d 1_{\{m(i)=m\}} (1 - c_{id}) c_{id} f_{id} w_d^2 s_{ijd} \left(\sum_{k \in \mathcal{N}_m \cup \{0\}} p_{mk} s_{ikd} - p_{mj} \right)$$

- The second term accounts for the effect of consumer choices on payments from MCOs to hospitals. As the price of hospital j rises, consumers will switch to cheaper hospitals.

Model

- Different implied behaviors according to coinsurance rates
- Increasing price:
  - decreases patients' expected utility (i.e. decreases MCOs' surplus)
  - might divert him to a lower/higher priced alternative (i.e. ambiguous impact on MCOs' costs)
- Platform effect on price levels is ambiguous: effective price sensitivity $\Omega + \Lambda$ may be smaller or higher than actual price sensitivity
- Mergers generally lead to higher prices (less competition between hospitals). Bargaining can amplify or mitigate this effect.

Estimation

Estimation of the patient choice stage

Multinomial logit estimation is carried out by maximum likelihood:

- Data includes information on patient residences, prices paid by MCOs to hospitals and coinsurance rates.
- Outcome variable is the choice of hospital.
- Estimated parameters: β and α using choice probability for patient i with disease d

$$s_{ijd}\left(\mathcal{N}_{m(i)}, \mathbf{p}_{m(i)}\right) = \frac{\exp\left(\beta X_{ijd} - \alpha c_{id} w_d p_{m(i)j}\right)}{\sum_{k \in \mathcal{N}_{m(i)} \cup \{0\}} \exp\left(\beta X_{ikd} - \alpha c_{id} w_d p_{m(i)k}\right)}$$

Estimation

Estimation of the first stage

- Estimation is carried out by GMM.
- Estimated parameters: bargaining weights ($\mathbf{b}$), cost determinants (γ), weight MCOs put on WTP measure (τ).
- Marginal cost residuals obtained using the bargaining model at estimated demand parameters:

$$\varepsilon(\mathbf{b}, \gamma, \tau) = -\gamma \mathbf{v} + \mathbf{mc}(\mathbf{b}, \tau) = -\gamma \mathbf{v} + \mathbf{p} + (\Omega + \Lambda(\mathbf{b}, \tau))^{-1} \mathbf{q}$$

- Moments are constructed based on marginal cost residuals:

$$\mathbb{E}[\varepsilon(\mathbf{b}, \gamma, \tau) \mid \mathbf{Z}] = 0$$

- Exogenous variables $\mathbf{Z}$ include:
  - Cost fixed effects $\mathbf{v}$.
  - Predicted (at overall mean price) WTP for the hospital, WTP for the system, WTP per enrollee for each MCO
  - Predicted total hospital quantity $\mathbf{q}$.

Results

Multinomial logit demand estimates

Variable	Coefficient	Standard error
Base price × weight × coinsurance	−0.0008**	(0.0001)
Travel time	−0.1150**	(0.0026)
Travel time squared	−0.0002**	(0.0000)
Closest	0.2845**	(0.0114)
Travel time × beds / 100	−0.0118**	(0.0008)
Travel time × age / 100	−0.0441**	(0.0023)
Travel time × FP	0.0157**	(0.0011)
Travel time × teach	0.0280**	(0.0010)
Travel time × residents/beds	0.0006**	(0.0000)
Travel time × income / 1000	0.0002**	(0.0000)
Travel time × male	−0.0151**	(0.0007)
Travel time × age 60+	−0.0017	(0.0013)
Travel time × weight / 1000	11.4723**	(0.4125)
Cardiac major diagnostic class × cath lab	0.2036**	(0.0409)
Obstetric major diagnostic class × NICU	0.6187**	(0.0170)
Nerv, circ, musc major diagnostic classes × MRI	−0.1409**	(0.0460)

Results

Demand elasticity estimates

Hospital	(1) PW	(2) Fairfax	(3) Reston	(4) Loudoun	(5) Fauquier
1. Prince William	−0.125	0.052	0.012	0.004	0.012
2. Inova Fairfax	0.011	−0.141	0.018	0.006	0.004
3. HCA Reston	0.008	0.055	−0.149	0.022	0.002
4. Inova Loudoun	0.004	0.032	0.037	−0.098	0.001
5. Fauquier	0.026	0.041	0.006	0.002	−0.153
6. Outside option	0.025	0.090	0.022	0.023	0.050

Note: Elasticity is $\frac{\partial s_j}{\partial p_k} \frac{p_k}{s_j}$ where j denotes row and k denotes column)

Results

Bargaining parameter estimates

	Specification 1		Specification 2	
Parameter	Estimate	S.E.	Estimate	S.E.
MCO welfare weight (τ)	2.79	(2.87)	6.69	(5.53)
MCO 1 bargaining weight	0.5	–	0.52	(0.09)
MCOs 2 & 3 bargaining weight	0.5	–	1.00**	(7.77×10^{-10})
MCO 4 bargaining weight	0.5	–	0.76**	(0.09)
Hospital cost parameters				
Prince William Hospital	8,635**	(3,009)	5,971**	(1,236)
Inova Alexandria	10,412*	(4,415)	6,487**	(1,905)
Inova Fairfax	10,786**	(3,765)	6,133**	(1,211)
Inova Fair Oaks	11,192**	(3,239)	6,970**	(2,352)
Inova Loudoun	12,014**	(3,188)	8,167**	(1,145)
Inova Mount Vernon	10,294*	(5,170)	4,658	(3,412)
Fauquier Hospital	14,553**	(3,390)	9,041**	(1,905)
No. VA Community Hosp.	10,086**	(2,413)	5,754**	(2,162)
Potomac Hospital	11,459**	(2,703)	7,653**	(902)
Reston Hospital Center	8,249**	(3,064)	5,756**	(1,607)
Virginia Hospital Center	7,993**	(2,139)	5,303**	(1,226)

Results

Counterfactual merger

Counterfactual	System	%Δ Price	%Δ Quantity	%Δ Profits
1. Inova/PWH merger	Inova & PWH	3.1	−0.5	9.3
	Rival hospitals	3.6	1.2	12.0
	Change at Inova+PW relative to PW base	30.5	−4.9	91.5
2. Inova/PWH merger with separate bargaining	Inova & PWH	3.3	−0.5	8.8
	Rival hospitals	3.5	1.2	11.2

Results

Counterfactual coinsurance levels

Counterfactual	System	%Δ Price	%Δ Quantity	%Δ Profits
1. No coinsurance	All hospitals	3.7	0.01	9.8
2. Coinsurance 10 times current	All hospitals	−16.1	0.9	−0.4
3. Inova/PWH merger, no coinsurance	Inova & PWH	2.9	0	7.4
	Rival hospitals	1.3	0	3.9

Consumer Welfare

- A common use of empirical demand models is to compute consumer welfare
- Compensating variation measures change in a consumer's income that equates utility in a particular economic environment to some chosen benchmark utility (Hicks, 1946).
- The methods can be used broadly for evaluating:
    - mergers, regulation
    - Consumer Price Index
- For example to measure consumer surplus change after a merger in discrete choice models
- We will focus on welfare gains from the introduction of new goods

Consumer Welfare Using Discrete Choice Model

- Assume the indirect utility is given by

$$u_{ijt} = x_{jt}\beta_i + \alpha_i p_{jt} + \xi_{jt} + \varepsilon_{ijt}$$

- With ε_{ijt} i.i.d. extreme value, the *inclusive value* (or social surplus) from a subset $A \subseteq \{1, 2, .., J\}$ of alternatives:

$$\begin{aligned} \omega_{iAt} &= E_{\varepsilon_{ijt}}\left(\max_{j\in A} u_{ijt}\right) \\ &= \ln\left(\sum_{j\in A} \exp\left\{x_{jt}\beta_i - \alpha_i p_{jt} + \xi_{jt}\right\}\right) \end{aligned}$$

- The expected utility from A prior to observing $(\varepsilon_{i0t}, ...\varepsilon_{iJt})$, knowing choice will maximize utility after observing shocks.
  - If no heterogeneity ($\beta_i = \beta$, $\alpha_i = \alpha$), captures average utility
  - With heterogeneity, need to integrate over it
  - If utility linear in price convert to dollars by dividing by α_i
  - With income effects conversion to dollars done by simulation

Red-bus-Blue-bus Example

- Originally, used to show the IIA problem of Logit
- Worst case scenario for Logit
- Consumers choose between driving car to work or (red) bus
  - working at home not an option
  - decision of whether to work does not depend on transportation
- Half the consumers choose the car and half choose the red bus
- Artificially introduce a new option: a blue bus
  - consumers color blind
  - no price or service changes
- In reality half the consumers choose car, rest split between the two color buses
- Consumer welfare has not changed

Red-bus-Blue-bus Example

- Suppose we want to use the Logit model to analyze consumer welfare generated by the introduction of the blue bus

$$u_{ijt} = \xi_{jt} + \varepsilon_{ijt}$$

- Normalizing $\xi_{car0} = 0$, therefore $\xi_{bus0} = 0$ because same market share
- Welfare at $t = 0$

$$\begin{aligned}\omega_{i0} &= E_{\varepsilon_{ij0}}\left(\max_j u_{ij0}\right) \\ &= \ln\left(\exp \xi_{00} + \exp \xi_{01}\right)\end{aligned}$$

Red-bus-Blue-bus Example

$$u_{ijt} = \xi_{jt} + \varepsilon_{ijt}$$

$t = 0$			$t = 1$			
	observed		predicted		observed	
option	share	ξ_{j0}	share	ξ_{j1}	share	ξ_{j1}
car	0.5					
red bus	0.5					
blue bus	–					
welfare						

Red-bus-Blue-bus Example

$$u_{ijt} = \xi_{jt} + \varepsilon_{ijt}$$

$t = 0$			$t = 1$			
	observed		predicted		observed	
option	share	ξ_{j0}	share	ξ_{j1}	share	ξ_{j1}
car	0.5	0				
red bus	0.5	0				
blue bus	–	–				
welfare	$\ln(2)$					

Red-bus-Blue-bus Example

- Let's introduce the blue bus
- If nothing changed, one might be tempted to hold ξ_{jt} fixed that is

$$\xi_{j1} = \xi_{j0} = 0$$

Red-bus-Blue-bus Example

$$u_{ijt} = \xi_{jt} + \varepsilon_{ijt}$$

$t = 0$			$t = 1$			
	observed		predicted		observed	
option	share	ξ_{j0}	share	ξ_{j1}	share	ξ_{j1}
car	0.5	0	0.33	0		
red bus	0.5	0	0.33	0		
blue bus	–	–	0.33	0		
welfare	$\ln(2)$		$\ln(3)$			

Red-bus-Blue-bus Example

- We obtained the usual result: with predicted shares Logit gives welfare gains
- Now, suppose we observed actual shares

Red-bus-Blue-bus Example

$$u_{ijt} = \xi_{jt} + \varepsilon_{ijt}$$

$t=0$			$t=1$			
	observed		predicted		observed	
option	share	ξ_{j0}	share	ξ_{j1}	share	ξ_{j1}
car	0.5	0	0.33	0	0.5	
red bus	0.5	0	0.33	0	0.25	
blue bus	–	–	0.33	0	0.25	
welfare	$\ln(2)$		$\ln(3)$			

Red-bus-Blue-bus Example

- To rationalize observed shares we need to let ξ_{jt} vary
- What did it exactly mean to introduce blue bus?

Red-bus-Blue-bus Example

$$u_{ijt} = \xi_{jt} + \varepsilon_{ijt}$$

$t = 0$			$t = 1$			
	observed		predicted		observed	
option	share	ξ_{j0}	share	ξ_{j1}	share	ξ_{j1}
car	0.5	0	0.33	0	0.5	0
red bus	0.5	0	0.33	0	0.25	$\ln(0.5)$
blue bus	–	–	0.33	0	0.25	$\ln(0.5)$
welfare	$\ln(2)$		$\ln(3)$		$\ln(2)$	

Generalizing from the example

- In the example, the Logit model fails in the first step (counterfactual simulation)
- Holds more generally:
  - with Logit, expected utility is $\ln\left(\sum_j \exp\left\{x_{jt}\beta - \alpha p_{jt} + \xi_{jt}\right\}\right) = \ln(1/s_{0t})$
  - since s_{0t} did not change in the observed data the Logit model predicted no welfare gain

Generalizing from the example

- With heterogeneity, Logit is wrong for evaluation of compensating variation
  - Logit change in expected utility

$$\ln\left(\frac{1}{s_{0,t}}\right) - \ln\left(\frac{1}{s_{0,t-1}}\right) = \ln\left(\frac{s_{0,t-1}}{s_{0,t}}\right) = \ln\left(\frac{\int s_{i,0,t-1}\, dP_\tau(\tau)}{\int s_{i,0,t}\, dP_\tau(\tau)}\right)$$

  and with random coefficients

$$\int \left[\ln\left(\frac{1}{s_{i,0,t}}\right) - \ln\left(\frac{1}{s_{i,0,t-1}}\right)\right] dP_\tau(\tau) = \int \ln\left(\frac{s_{i,0,t-1}}{s_{i,0,t}}\right) dP_\tau(\tau)$$

  - Difference depends on the change in the heterogeneity in the probability of choosing the outside option, $s_{i,0,t}$
  - Difference can be positive or negative

Generalizing from the example

- The key in the above example is that ξ_{jt} was allowed to change to fit the data.
- This works when we see data pre and post (allows us to tell how we should change ξ_{jt})
- What if we do not have data for the counterfactual?
    - have a model of how ξ_{jt} is determined
    - make an assumption about how ξ_{jt} changes
    - bound the effects

Price index and quantity index

- Nevo (ReStat, 2003) uses the latter approach to compute price indexes based on estimated demand systems
- Price index:
  - Summarize many prices into a single index which tries to measure the average price increase.
  - Different weights to average across products will yield different indices. Quantities consumed in periods t or $t+1$ are common choices.
- Quantity index:
  - For example, given prices in $t+1$, is the bundle consumed in t affordable? If so, and if it was not chosen, then consumer is better off (in terms of revealed preference).
  - Can use the prices in $t+1$ to average across quantities in both periods to construct a (revealed preference) quantity index. As for price index, there are many choices for the weights to average quantities.

Vertical Relations

- Bernheim and Whinston (1985)
- Villas-Boas (2007)
- Bonnet and Dubois (2010, 2015)
- Rey and Vergé (2010)
- ..

Bonnet and Dubois “Inference on Vertical Contracts between Manufacturers and Retailers Allowing for Nonlinear Pricing and Resale Price Maintenance” (2010)

- Modelling intermediary strategic behavior between producer and consumer
- Random Coefficient Logit

$$V_{ijt} = \beta_j + \gamma_t - \alpha_i p_{jt} + \xi_{jt} + \varepsilon_{ijt}$$

with

$$\alpha_i = \alpha + \sigma v_i$$

- Estimation of random coefficients logit model on aggregate data
- Instruments: price of inputs (wage salary index, diesel fuel and packaging material price indices, interacted with dummy variables on characteristics of each product)
- Test robustness to aggregation bias (aggregation by month and from different distant consumers)

Demand Results

TABLE 4 Estimation Results of Demand Models

Coefficients (Standard Error)	ML (1)	RCL (2)	RCL (Robustness Check) (3)
Price ($-\alpha$)	−5.47 (0.44)	−8.95 (1.14)	−10.74 (1.45)
Price (σ)		2.04 (0.81)	3.61 (1.20)
Standard deviation of price			0.81 (1.13)
Average distance			0.03 (0.06)
Coefficients β_j, γ_t not shown			
Overidentifying restrictions test	6.30 ($\chi^2(10)$)	7.81 ($\chi^2(3)$)	12.50 ($\chi^2(8)$)

Supply side model - Linear Prices

- Taking into account the strategic role of vertical intermediaries
- Retailers profits

$$\Pi^r = \sum_{j \in S_r} (p_j - w_j - c_j) s_j(\mathbf{p}) M$$

- Bertrand-Nash equilibrium in prices

$$s_j + \sum_{k \in S_r} (p_k - w_k - c_k) \frac{\partial s_k}{\partial p_j} = 0, \quad \text{for all } j \in S_r$$

- Profit of manufacturer f is

$$\Pi^f = \sum_{j \in G_f} (w_j - \mu_j) s_j(p(w)) M$$

- Pure-strategy Bertrand-Nash equilibrium

$$s_j + \sum_{k \in G_f} \sum_{l=1,..,J} (w_k - \mu_k) \frac{\partial s_k}{\partial p_l} \frac{\partial p_l}{\partial w_j} = 0, \quad \text{for all } j \in G_f$$

Derivatives of retail prices with respect to wholesale prices by total differentiation of retailer's first order conditions

Supply side model - Non Linear Prices

- Two-part Tariffs (Rey and Vergé, 2004)

$$\Pi^r = \sum_{j \in S_r} [M(p_j - w_j - c_j)s_j(\mathbf{p}) - F_j]$$

where F_j is franchise fee.

- Manufacturers set w_k and F_k to maximize

$$\Pi^f = \sum_{k \in G_f} [M(w_k - \mu_k)s_k(\mathbf{p}) + F_k]$$

subject to constraints $\Pi^r \geq \overline{\Pi}^r$, for all $r = 1, .., R$

- Rey and Vergé (2004): participation constraints must be binding:

$$\Pi^f = \sum_{k \in G_f} (p_k - \mu_k - c_k)s_k(\mathbf{p}) + \sum_{k \notin G_f} (p_k - w_k - c_k)s_k(\mathbf{p}) - \sum_{j \notin G_f} F_j$$

Supply side model - Non Linear Prices

- Binding participation constraints ($M = 1, \overline{\Pi}^r = 0$):

$$\sum_{j \in S_r} F_j = \sum_{j \in S_r} [(p_j - w_j - c_j) s_j(\mathbf{p})]$$

$$\begin{aligned}
\Pi^f &= \sum_{k \in G_f} [(w_k - \mu_k) s_k(\mathbf{p})] + \sum_{k \in G_f} F_k \\
&= \sum_{k \in G_f} [(w_k - \mu_k) s_k(\mathbf{p})] + \sum_k F_k - \sum_{k \notin G_f} F_k \\
&= \sum_{k \in G_f} [(w_k - \mu_k) s_k(\mathbf{p})] + \sum_r \left(\sum_{k \in S_r} F_k \right) - \sum_{k \notin G_f} F_k \\
&= \sum_{k \in G_f} (w_k \text{-} \mu_k) s_k(\mathbf{p}) + \sum_r [\sum_{j \in S_r} (p_j - w_j - c_j) s_j(\mathbf{p})] - \sum_{k \notin G_f} F_k \\
&= \sum_{k \in G_f} [(w_k - \mu_k) s_k(\mathbf{p})] + \sum_k (p_k - w_k - c_k) s_k(\mathbf{p}) - \sum_{k \notin G_f} F_k \\
&= \sum_{k \in G_f} (p_k - \mu_k - c_k) s_k(\mathbf{p}) + \sum_{k \notin G_f} (p_k - w_k - c_k) s_k(\mathbf{p}) - \sum_{j \notin G_f} F_j
\end{aligned}$$

Supply side model - Non Linear Prices

- Resale Price Maintenance (RPM) equilibrium :

$$\max_{\{p_k,w_k,F_k\}\in G_f} \Pi^f = \max_{\{p_k\}\in G_f} \Pi^f$$

- RPM equilibrium $p^*(w^*)$ satisfies for all f

$$\max_{\{p_k\}\in G_f} \sum_{k\in G_f} (p_k - \mu_k - c_k) s_k(\mathbf{p}) + \sum_{k\notin G_f} (p_k^* - w_k^* - c_k) s_k(\mathbf{p})$$

- Thus, first order conditions are

$$\sum_{k\in G_f} (p_k - \mu_k - c_k) \frac{\partial s_k(\mathbf{p})}{\partial p_j} + s_j(\mathbf{p}) + \sum_{k\notin G_f} (p_k^* - w_k^* - c_k) \frac{\partial s_k(\mathbf{p})}{\partial p_j} = 0$$

Supply side model - Non Linear Prices - RPM

- Multiple equilibria with RPM:
  - $w_k^* = \mu_k$: retailers as residual claimants and manufacturers capture full monopoly rents through fixed fees. FOC are

$$\sum_{k=1}^{J} (p_k - \mu_k - c_k) \frac{\partial s_k(\mathbf{p})}{\partial p_j} + s_j(\mathbf{p}) = 0$$

  - $p_k^*(\mathbf{w}^*) - w_k^* - c_k = 0$:

$$\sum_{k \in G_f} (p_k - \mu_k - c_k) \frac{\partial s_k(\mathbf{p})}{\partial p_j} + s_j(\mathbf{p}) = 0$$

- Pricing decisions implemented by manufacturers and share of profits between retailers and manufacturers unidentified

Supply side model - Non Linear Prices - Without RPM

- Without RPM

$$\max_{\{w_k\}\in G_f} \sum_{k\in G_f} (p_k - \mu_k - c_k) s_k(\mathbf{p}) + \sum_{k\notin G_f} (p_k - w_k - c_k) s_k(\mathbf{p}).$$

Then, the first order conditions are

$$\begin{aligned} 0 \quad &= \quad \sum_k \frac{\partial p_k}{\partial w_i} s_k(\mathbf{p}) + \sum_{k\in G_f} \left[(p_k - \mu_k - c_k) \sum_j \frac{\partial s_k}{\partial p_j} \frac{\partial p_j}{\partial w_i} \right] \\ &\quad + \sum_{k\notin G_f} \left[(p_k - w_k - c_k) \sum_j \frac{\partial s_k}{\partial p_j} \frac{\partial p_j}{\partial w_i} \right] \end{aligned}$$

- Each supply side model is identified.

Inference on Supply Side Models

- Which supply side model is best?
- Testing between models using cost restrictions
- Considering model h, denote γ_{jt}^{h} retailer price cost margin and Γ_{jt}^{h} manufacturer price cost margin
- Marginal cost of production and distribution $C_{jt}^{h} = \mu_{jt}^{h} + c_{jt}^{h}$, such that

$$C_{jt}^{h} = p_{jt} - \Gamma_{jt}^{h} - \gamma_{jt}^{h}.$$

- Assume

$$\ln C_{jt}^{h} = \omega_{j}^{h} + W_{jt}^{\prime}\lambda_{h} + \ln \eta_{jt}^{h}$$

  and

$$E(\ln \eta_{jt}^{h} | \omega_{j}^{h}, W_{jt})$$

- Non-nested tests (Vuong, 1989, and Rivers and Vuong, 2002) are then applied to infer which model h is statistically the best.
- Idea: infer which cost equation has best statistical fit given the observed cost shifters W_{jt}

Inference on Supply Side Models

- Test each model against each other, for models h and h'

$$p_{jt} = \Gamma_{jt}^{h} + \gamma_{jt}^{h} + \left[\exp(\omega_j^h + W'_{jt}\lambda_h)\right] \eta_{jt}^{h}$$

and

$$p_{jt} = \Gamma_{jt}^{h'} + \gamma_{jt}^{h'} + \left[\exp(\omega_j^{h'} + W'_{jt}\lambda_{h'})\right] \eta_{jt}^{h'}.$$

Using NLLS

$$\begin{aligned}
\min_{\lambda_h,\omega_j^h} Q_n^h(\lambda_h,\omega_j^h) &= \min_{\lambda_h,\omega_j^h} \frac{1}{n}\sum\nolimits_{j,t}\left(\ln \eta_{jt}^h\right)^2 \\
&= \min_{\lambda_h,\omega_j^h} \frac{1}{n}\sum\nolimits_{j,t}\left[\ln\left(p_{jt}-\Gamma_{jt}^h-\gamma_{jt}^h\right)-\omega_j^h - W'_{jt}\lambda_h\right]
\end{aligned}$$

Inference on Supply Side Models

- Null hypothesis: non-nested models *asymptotically equivalent* when

$$H_0 \;:\; \lim_{n\to\infty}\left\{\bar{Q}_n^h(\overline{\lambda}_h,\overline{\omega}_j^h)-\bar{Q}_n^{h'}(\overline{\lambda}_{h'},\overline{\omega}_j^{h'})\right\}=0$$

where $\bar{Q}_n^h(\overline{\lambda}_h,\overline{\omega}_j^h)$ (resp. $\bar{Q}_n^{h'}(\overline{\lambda}_{h'},\overline{\omega}_j^{h'})$) are expectation of a lack-of-fit criterion $Q_n^h(\lambda_h,\omega_j^h)$

- Alternative hypothesis: h is *asymptotically better* than h' when

$$H_1 \;:\; \lim_{n\to\infty}\left\{\bar{Q}_n^h(\overline{\lambda}_h,\overline{\omega}_j^h)-\bar{Q}_n^{h'}(\overline{\lambda}_{h'},\overline{\omega}_j^{h'})\right\}<0.$$

h' is *asymptotically better* than h when

$$H_2 \;:\; \lim_{n\to\infty}\left\{\bar{Q}_n^h(\overline{\lambda}_h,\overline{\omega}_j^h)-\bar{Q}_n^{h'}(\overline{\lambda}_{h'},\overline{\omega}_j^{h'})\right\}>0.$$

Inference on Supply Side Models

- Test statistic $T_n = \frac{\sqrt{n}}{\hat{\sigma}_n^{hh'}} \left\{ Q_n^h(\widehat{\lambda}_h, \widehat{\omega}_j^h) - Q_n^{h'}(\widehat{\lambda}_{h'}, \widehat{\omega}_j^{h'}) \right\}$ captures statistical variation of sample values of the lack-of-fit criterion
- $\hat{\sigma}_n^{hh'}$ denotes the estimated value of the variance of the difference in lack-of-fit.
- Rivers and Vuong showed that T_n is standard normal.

Goldberg, 1995

- Goldberg, "Product Differentiation and Oligopoly in International Markets: The Case of the Automobile Industry" (1995)
- Endogeneity problem solved using household-level data
- Nested Logit versus RC Logit (many nests to avoid IIA as much as possible)
- Objective of the paper:
    - study imperfect competition in international trade, welfare impacts of different trade policies
        - In the 80's, U.S. imports of cars subject to a tariff rate of 2.9%
        - Japanese auto sales in the US limited by "Voluntary Export Restraint" (VER)

Demand Model

- Nested Logit nests determined by buy/not buy, new/used, country of origin (foreign vs. domestic) and segment

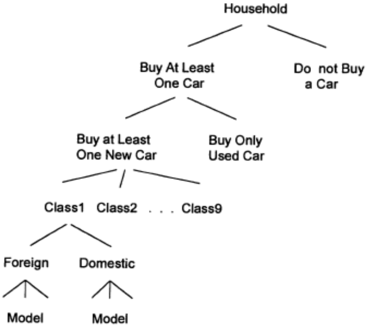

FIGURE 1.—Automobile choice model.

- This model can be viewed as using segment and country of origin as (dummy) characteristics, and assuming a particular distribution on their coefficients.

Data

- Household-level survey from the Consumer Expenditure Survey:
  - 20,571, households between 83-87
  - 6,172 (30%) bought a car
  - 1,992 (33%) new car
  - 1,394 (70%) domestic and 598 foreign
- Prices (and characteristics) are obtained from Automotive News Market Data Book

Estimation

- The model is estimated by ML
- The likelihood is partitioned and estimated recursively:
  - At the lowest level the choice of model conditional on origin, segment and newness, based on the estimated parameters, an "inclusive value" is computed and used to estimate the choice of origin conditional on segment and newness, etc.
- Does not deal with endogeneity. Origin and segment fixed effects are included, but these do not fully account for brand unobserved characteristics

Table II: Price elasticities by class

TABLE II

PRICE ELASTICITIES OF DEMAND (AVERAGE BY CLASS)

Class	Origin	Elasticity	Elasticity (first time buyer)	Elasticity (repeat buyer)
Subcompacts	DOM	−3.2857	−3.6245	−2.9816
	FOR	−3.6797	−5.2531	−2.9488
Compacts	DOM	−3.419	−4.8722	−3.1546
	FOR	−4.0319	−5.7229	−3.3733
Intermediate	DOM	−4.1799	−5.3153	−2.8420
	FOR	−5.1524	−6.2232	−4.9274
Standard	DOM	−4.7121	−5.932	−4.3730
Luxury	DOM	−1.9121	−2.5981	−1.1137
	FOR	−2.7448	−3.1272	−1.9959
Sports	DOM	−1.0654	−2.3468	−1.3959
	FOR	−1.5254	3.0211	−1.1429
Pick-ups	DOM	−3.5259	−5.1391	−3.1647
	FOR	−2.6883	−3.9822	−2.1483
Vans	DOM	−4.3633	−5.4977	−3.9790
	FOR	−4.6548	−4.8837	−2.4376
Other	DOM	−4.0884	−4.3185	−3.5694
	FOR	−3.0271	−3.3185	−2.3345

Table III: Price semi-elasticities

TABLE III

CROSS PRICE SEMI-ELASTICITIES FOR SELECTED MODELS

	Chevette	Civic	Tercel	Escort	Accord
Chevette	−3.2	49.1E − 07	16.4E − 07	0.9E − 07	9.0E − 07
Civic	7.6E − 07	−3.4	35.5E − 07	0.8E − 07	14.9E − 07
Tercel	7.7E − 07	109.8E − 07	−3.4	0.8E − 07	11.6E − 07
Escort	6.3E − 07	34.6E − 07	11.3E − 07	−3.4	12.5E − 07
Accord	6.1E − 07	66.2E − 07	16.2E − 07	1.3E − 07	−3.4
Mazda 626	6.4E − 07	50.1E − 07	15.3E − 07	1.7E − 07	72.2E − 07
Century	5.5E − 07	28.0E − 07	9.6E − 07	0.8E − 07	7.1E − 07
Skylark	5.5E − 07	28.6E − 07	9.9E − 07	0.8E − 07	7.1E − 07
Audi 5000	5.7E − 07	48.6E − 07	16.6E − 07	0.8E − 07	10.1E − 07
Diplomat	4.9E − 07	25.5E − 07	8.7E − 07	0.8E − 07	6.6E − 07
Cad. Fleetwood	0.3E − 07	2.1E − 07	0.7E − 07	0.1E − 07	0.5E − 07
Park Avenue	0.3E − 07	2.1E − 07	0.7E − 07	0.1E − 07	0.5E − 07
Jaguar	0.3E − 07	3.2E − 07	1.0E − 07	0.0E − 07	0.6E − 07
Fiero	0.4E − 07	3.0E − 07	1.0E − 07	0.1E − 07	0.7E − 07
Ferrari	0.4E − 07	4.0E − 07	1.3E − 07	0.1E − 07	0.8E − 07

Supply side with Quotas and Markups

- The profits of firm f, with wholesale price p_{it} (in \$US), marginal cost c_{jt} :

$$\Pi_{ft} = \sum_{j=1}^{n_{ft}} (e_t p_{jt} - c_{jt}) q_{jt}(\mathbf{p}_t)$$

where e_t is exchange rate and with quotas constraints for foreign firms

$$\sum_{j \in V_f} q_{jt} \leq D_{ft}$$

- Quotas can affect a subset of produced cars (passengers cars only for Japanese manufacturers)

Supply side with Quotas and Markups

- Then, the first order conditions are:
  - for domestic firms

$$q_{kt}(\mathbf{p}) + \sum_{j=1}^{n_{ft}} (p_{jt} - c_{jt}) \frac{\partial q_{jt}(\mathbf{p})}{\partial p_{kt}} = 0$$

  - for foreign firms

$$e_t q_{kt}(\mathbf{p}) + \sum_{j=1}^{n_{ft}} (e_t p_{jt} - c_{jt}) \frac{\partial q_{jt}(\mathbf{p})}{\partial p_{kt}} = 0$$

  - for foreign firms subject to quotas

$$e_t q_{kt}(\mathbf{p}) + \sum_{j=1}^{n_{ft}} (e_t p_{jt} - c_{jt}) \frac{\partial q_{jt}(\mathbf{p})}{\partial p_{kt}} - \lambda_{ft} \sum_{j \in V_f} \frac{\partial q_{jt}(\mathbf{p})}{\partial p_{kt}} = 0$$

with $\lambda_{ft} \geq 0$, $\sum_{j \in V_f} q_{jt} \leq D_{ft}$.

Supply side with Quotas and Markups

- As some manufacturers produce only passenger cars, all cars are subject to quota, then quota effect not separately identified from marginal cost
- Goldberg (1995) assumes $\lambda_{ft} = \lambda_t$ for identification. Corresponds to assumption that Japanese government allocated quotas to equalize shadow price of quantity constraint among firms.
- Not completely implausible: by allocating quotas on basis of previous year's market shares Japanese government was essentially trying to equalize "cost" of VER across firms.
- Then marginal costs and markups are identified.

Table IV: Implied Markups

Model	Cost	Price	Markup	(Price − Cost)
Civic	4884	5680	0.14	796
Escort	3068	4565	0.33	1497
Lynx	3069	4325	0.29	1256
Accord	5286	5854	0.10	567
Audi 5000	7353	14165	0.48	6812
Oldsmobile 98	5372	11295	0.52	5923
Jaguar	10768	19091	0.44	8323
Mercedes 300	13188	22662	0.42	9474
Porsche 944	5714	13136	0.56	7422
Ferrari	7679	19698	0.61	12018

Simulations

- Then, model is used to simulate counterfactuals:
  - Quantify the impact of quotas on sales, prices, and quality mix of domestic and foreign firms, by solving the model under two different assumptions, quotas vs. free trade
  - Equivalent tariff to VER:

$$\sum_{i=1}^{n_{ft}} \left(\frac{e_t p_{it}}{1+TR} - c_{it}\right) q_{it}(\mathbf{p}_t)$$

    solve for price and tariff levels that satisfy first order conditions of the profit maximizing producers and constraint that Japanese passenger car sales do not exceed the VER level.
  - Exchange Rate Pass-Through in the Automobile Industry

Simulations: Effects on Sales

- Quotas led to drop in Japanese sales of passenger cars
- Did not benefit American producers to the full extent (on average only 54%) because of:
  - Substitution towards other imports, mainly European cars (+18%)
  - Small decline in total sales (consequence of higher prices) because substitution towards used cars

Simulations: Effects on Sales

TABLE VI

EFFECTS OF THE VER ON SALES

Year	Total	Passenger Cars	American	Japanese	Other
1983	−33,478 (0.3%)	−90,899 (1.1%)	+209,990 (2.0%)	−340,321 (15.0%)	+39,432 (7.0%)
1984	−37,808 (0.3%)	−124,990 (1.6%)	+140,807 (1.5%)	−298,477 (9.4%)	+32,680 (6.4%)
1987	−29,991 (0.02%)	−106,632 (1.0%)	+57,510 (0.5%)	−185,843 (4.5%)	+21,702 (2.8%)

Simulations: Effects on Market Shares

- Decrease in Japanese subcompact sales larger than decrease of compact sales, implying the share of the relatively more expensive compacts in Japanese imports of passenger cars increases.
- Share of higher priced intermediate, standard, and luxury automobiles in American and foreign production rises. Overall "quality" of automobiles increases.
- Quotas change relative prices of market segments so that there is substitution towards higher priced cars.

Simulations: Effects on Market Shares

TABLE VII

PERCENT CHANGE IN MARKET SEGMENTS (1983)

Class	Domestic	Foreign
Subcompact	1.9	−18.0
Compact	2.5	−15.3
Intermediate	4.2	7.5
Standard	3.3	—
Luxury	3.0	4.4
Sports	4.2	7.6
Trucks	5.0	10.1
Vans	0.3	0.8
Other	3.1	4.0

Goldberg and Hellerstein (2013)

- Goldberg, P.K. and R. Hellerstein (2013) "A Structural Approach to Identifying the Sources of Local-Currency Price Stability", Review of Economic Studies, 80(1), 175-210.
- Inertia of the local currency prices of traded goods in the face of exchange rate changes is a well-documented phenomenon in International Economics.
- Develops structural model to identify the sources of local currency price stability (application on the beer market)
- Exploits manufacturers' and retailers' first-order conditions with information on frequency of price adjustments following exchange rate changes to quantify:
  - local non-traded cost components
  - markup adjustments
  - nominal price rigidities

Results

- On average, find that incomplete exchange rate pass-through is:
  - 60% due to local nontraded costs
  - 8% to markup adjustment
  - 30% to existence of own brand price adjustment costs
  - 1% to the indirect/strategic effect of such costs

Reduced form regressions

- Reduced forms equations for product j week t, zone z

$$\begin{aligned}
\ln p_{jzt}^{r} &= c_j + \zeta_z + \theta_t + \alpha \ln e_{jt} + \beta \ln co_{jt} + \varepsilon_{jzt} \\
\ln p_{jzt}^{w} &= c_j + \zeta_z + \theta_t + \alpha \ln e_{jt} + \beta \ln co_{jt} + \varepsilon_{jzt} \\
\ln p_{jzt}^{r} &= c_j + \zeta_z + \theta_t + \alpha \ln p_{jzt}^{w} + \varepsilon_{jzt}
\end{aligned}$$

  where p_{jzt}^{r} is retail price; p_{jzt}^{w} wholesale price; e_{jt} is bilateral nominal exchange rate (domestic currency units per unit of foreign currency); co_{jt} are proxy for cost shocks (domestic (U.S.) wages, price of barley in each country producing beer, price of electricity in Chicago area, wages in each beer exporting country)
- Variables in levels:
  - focus on long-run pass-through of exchange rate changes, and not short-term dynamics.

Reduced form regressions

	Retail price	Retail price	Wholesale price	Wholesale price	Retail price
Exchange rate	5·96 (1·50)**	6·72 (1·56)**	4·27 (1·50)**	4·74 (1·52)**	
Wholesale price					105·37 (2·53)**
R^2	0·65	0·65	0·81	0·81	0·80

Notes: The dependent variable is the retail or the wholesale price for a six-pack of each brand of beer. The exchange rate is the average of the previous week's bilateral spot rate between the foreign manufacturer's country and the U.S. (dollars per unit of foreign currency). Includes brand, price zone, and week fixed effects. The second and fourth columns report results with controls for domestic and foreign costs. Robust standard errors in parentheses, those starred significant at the *5% or **1% level. 3636 observations. *Source:* Authors' calculations.

Prices

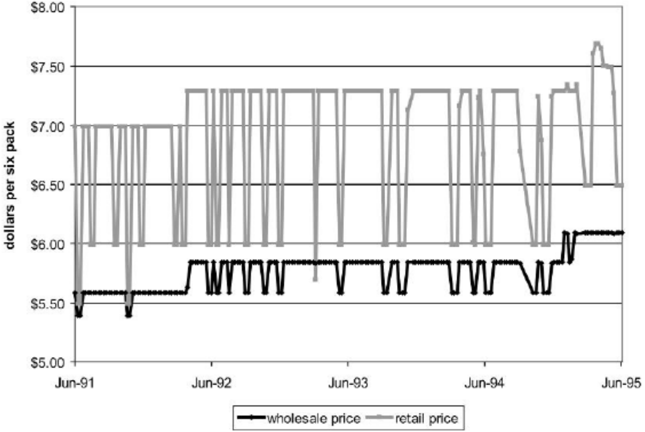

FIGURE 1

Weekly retail and wholesale prices for Bass Ale.

Prices are for a single six-pack and are from Zone 1. 202 observations. Source: *Dominick's*

Theory

- Retailers profits with fixed costs of changing prices A_{jt}^r

$$\Pi_t^r = \sum_j (p_{jt}^r - p_{jt}^w - ntc_{jt}^r) s_{jt}(\mathbf{p}_t^r) - A_{jt}^r$$

with

$$\begin{aligned} A_{jt}^r &= 0 \text{ if } p_{jt}^r = p_{jt-1}^r \\ A_{jt}^r &> 0 \text{ if } p_{jt}^r \neq p_{jt-1}^r \end{aligned}$$

Theory

- Case 1: $p^r_{jt} \neq p^r_{jt-1}$
    - Bertrand-Nash equilibrium in prices

$$s_{jt} + \sum_k (p^r_{kt} - p^w_{kt} - ntc^r_{kt}) \frac{\partial s_{kt}}{\partial p^r_{jt}} = 0$$

gives

$$\mathbf{p}^r_t = \mathbf{p}^w_t + \mathbf{ntc}^r_t - \Omega^{-1}_{rt} \mathbf{s}_t$$

- If retailer changed prices, extra profit associated to new price must be as large as adjustment cost: for all $j \neq k$

$$\begin{aligned} & \sum_k (p^r_{kt} - p^w_{kt} - ntc^r_{kt}) s_{kt} \left(\mathbf{p}^r_t\right) - A^r_{jt} \\ \geq \quad & (p^r_{jt-1} - p^w_{jt} - ntc^r_{jt}) s_{jt} (p^r_{jt-1}, \mathbf{p}^r_{-jt}) \\ & + \sum_{k \neq j} (p^r_{kt} - p^w_{kt} - ntc^r_{kt}) s^c_{kt} (p^r_{jt-1}, \mathbf{p}^r_{-jt}) \end{aligned}$$

where $s^c_{jt}(p^r_{jt-1}, \mathbf{p}^r_{-jt})$: counterfactual market share if price unchanged

Theory

- Then (revealed preference argument):

$$
\begin{aligned}
A_{jt}^{r} \quad \leq \quad \overline{A_{jt}^{r}} &= (p_{jt}^{r}\text{-}p_{jt}^{w}\text{-}ntc_{jt}^{r})s_{jt}(\mathbf{p}_{t}^{r}) - (p_{jt-1}^{r}\text{-}p_{jt}^{w}\text{-}ntc_{jt}^{r})s_{jt}(p_{jt-1}^{r}, \mathbf{p}_{-jt}^{r}) \\
&+ \sum\nolimits_{k\neq j}(p_{kt}^{r} - p_{kt}^{w} - ntc_{kt}^{r})\left[s_{kt}\left(\mathbf{p}_{t}^{r}\right) - s_{kt}^{c}(p_{jt-1}^{r}, \mathbf{p}_{-jt}^{r})\right]
\end{aligned}
$$

Theory

- Case 2: $p^r_{jt} = p^r_{jt-1}$

$$
\begin{aligned}
& (p^r_{jt-1} - p^w_{jt} - ntc^r_{jt}) s_{jt}(p^r_{jt-1}, \mathbf{p}^r_{-jt}) \\
& + \sum\nolimits_{k \neq j} (p^r_{kt} - p^w_{kt} - ntc^r_{kt}) s_{kt}(\mathbf{p}^r_t) \\
\geq \; & (p^{rc}_{jt} - p^w_{jt} - ntc^r_{jt}) s^c_{jt}(p^{rc}_{jt-1}, \mathbf{p}^r_{-jt}) \\
& + \sum\nolimits_{k \neq j} (p^r_{kt} - p^w_{kt} - ntc^r_{kt}) s^c_{jt}(p^{rc}_{jt}, \mathbf{p}^r_{-jt}) - A^r_{jt}
\end{aligned}
$$

where p^{rc}_{jt} denotes the counterfactual price the retailer would charge had he behaved according to optimality conditions and $s^c_{jt}(p^{rc}_{jt-1}, \mathbf{p}^r_{-jt})$ the counterfactual market share corresponding to optimal price.

Theory

- Then

$$\begin{aligned} A^r_{jt} &\geq \underline{A^r_{jt}} \\ &= (p^{rc}_{jt}\text{-}p^w_{jt}\text{-}ntc^r_{jt})s^c_{jt}(p^{rc}_{jt-1}, \mathbf{p}^r_{-jt}) - (p^r_{jt-1}\text{-}p^w_{jt}\text{-}ntc^r_{jt})s_{jt}(p^r_{jt-1}, \mathbf{p}^r_{-jt}) \\ &\quad \textstyle\sum_{k\neq j}(p^r_{kt} - p^w_{kt} - ntc^r_{kt}) \left[s^c_{jt}(p^{rc}_{jt}, \mathbf{p}^r_{-jt}) - s_{kt}\left(\mathbf{p}^r_t\right)\right] \end{aligned}$$

Theory

- Then, quantify adjustment costs as follows:
    - Estimate demand function.
    - Use first-order conditions for each product j to estimate nontraded costs and markups using only periods in which price p_{jt}^r adjusts
    - In periods when the price does not adjust, non-traded costs are not identified but can derive estimates by imposing additional structure (modelling non-traded costs parametrically as function of observables)
    - Compute counterfactual price p_{jt}^{rc} if there were no price rigidities and retailer would have behaved according to profit maximization
    - Compute counterfactual market share $s_{jt}^c(p_{jt-1}^{rc}, \mathbf{p}_{-jt}^r)$
    - Exploit inequalities to derive upper and lower bounds to adjustment costs A_{jt}^r.

Theory

- Profit of manufacturer f is

$$\Pi_t^w = \sum_{j \in S_w} (p_{jt}^w - c_{jt}^w \left(tc_{jt}^w, ntc_{jt}^w\right)) s_{jt}(\mathbf{p}_t^r(\mathbf{p}_t^w)) - A_{jt}^w$$

where A_{jt}^w denotes the price adjustment cost incurred by the manufacturer

$$\begin{aligned} A_{jt}^w &= 0 \text{ if } p_{jt}^w = p_{jt-1}^w \\ A_{jt}^w &> 0 \text{ if } p_{jt}^w \neq p_{jt-1}^w \end{aligned}$$

Theory

- Case 1: $p^w_{jt} \neq p^w_{jt-1}$
  - profit maximization of manufacturers

$$s_{jt} + \sum_{k \in S_w} (p^w_{kt} - c^w_{kt}) \frac{\partial s_{kt}}{\partial p^w_{jt}} = 0$$

gives

$$\mathbf{p}^w_t = \mathbf{c}^w_t - \Omega^{-1}_{wt} \mathbf{s}_t$$

where $\Omega_{wt} = \Omega'_{pt} \Omega_{rt}$ and Ω_{pt} is the matrix of partial derivative of each retail price with respect to each product's wholesale price

Theory

- If manufacturers changed prices, extra profit associated to new price must be as large as adjustment cost:

$$\begin{aligned} & \sum\nolimits_{k \in S_w} (p^w_{kt} - c^w_{kt}) s_{kt} \left(\mathbf{p}^r_t \left(\mathbf{p}^w_t \right) \right) - A^w_{jt} \\ \geq \; & (p^w_{jt-1} - c^w_{jt}) s^c_{jt} (p^{rc}_{jt}, \mathbf{p}^{rc}_{-jt}) + \sum\nolimits_{k \in S_w, k \neq j} (p^w_{kt} - c^w_{kt}) s^c_{kt} (p^{rc}_{jt}, \mathbf{p}^r_{-jt}) \end{aligned}$$

  where p^{rc}_{jt} is the counterfactual price the retailer would charge when faced with unchanged wholesale price p^w_{jt-1} as well as counterfactual retail prices $\mathbf{p}^{rc}_{-jt}$ of the other products offered by the retailer
- Problem:
  - given existence of retailer adjustment costs, counterfactual retail prices may or may not change from previous period.
  - To assess manufacturer's beliefs about the retailer's behavior: need retailer adjustment costs. However, we derive only bounds and not point estimates of these costs.

Theory

- Conjecture (not based on an equilibrium notion):
  - As retail prices empirically never adjust when wholesale price does not
  - Assume manufacturers believe the retail price will not change if does not adjust the wholesale price. Thus, assume

$$p_{jt}^{rc} = p_{jt-1}^{r} \text{ if } p_{jt}^{w} = p_{jt-1}^{w}$$

  - For counterfactual prices of other products, $\mathbf{p}_{-jt}^{rc}$, rely on observation that empirically retail price almost never changes when only wholesale prices of other products change, but only changes if wholesale price of same product has changed from previous period
- Then

$$A_{jt}^{w} \leq \overline{A_{jt}^{w}} = (p_{jt}^{w}\text{-}c_{jt}^{w})s_{jt}(\mathbf{p}_t^r(\mathbf{p}_t^w)) - (p_{jt-1}^{w}\text{-}c_{jt}^{w})s_{jt}(p_{jt}^{rc}, \mathbf{p}_{-jt}^{rc}) + \sum\nolimits_{k\in S_w, k\neq j}(p_{kt}^{w} - c_{kt}^{w})\left[s_{kt}(\mathbf{p}_t^r(\mathbf{p}_t^w)) - s_{kt}^{c}(p_{jt}^{rc}, \mathbf{p}_{-jt}^{rc})\right]$$

Theory

- Case 2: wholesale price does not change $p^w_{jt} = p^w_{jt-1}$
- It is possible that retailer adjusts price in periods when wholesale price does not change. In practice, do not observe this case, hence, concentrate on case where both wholesale and retail prices remain unchanged, i.e. $p^w_{jt} = p^w_{jt-1}$ and $p^r_{jt} = p^r_{jt-1}$.
- Then

$$\begin{aligned} & \sum\nolimits_{k \in S_w} (p^w_{kt} - c^w_{kt}) s_{kt} \left(p^r_{jt-1}, \mathbf{p}^r_{-jt} \right) \\ \geq \quad & (p^{wc}_{jt} - c^w_{jt}) s^c_{jt} (p^{rc}_{jt}, \mathbf{p}^{rc}_{-jt}) + \sum_{k \in S_w, k \neq j} (p^w_{kt} - c^w_{kt}) s^c_{kt} (p^{rc}_{jt}, \mathbf{p}^r_{-jt}) - A^w_{jt} \end{aligned}$$

and

$$\begin{aligned} A^w_{jt} \quad \geq \quad & \underline{A^w_{jt}} = (p^{wc}_{jt} \text{-} c^w_{jt}) s^c_{jt} (p^{rc}_{jt}, \mathbf{p}^{rc}_{-jt}) - (p^w_{jt-1} \text{-} c^w_{jt}) s_{jt} (p^r_{jt-1}, \mathbf{p}^r_{-jt}) \\ & + \sum\nolimits_{k \in S_w, k \neq j} (p^w_{kt} - c^w_{kt}) \left[s^c_{kt} \left(p^{rc}_{jt}, \mathbf{p}^{rc}_{-jt} \right) - s_{kt} (p^r_{jt-1}, \mathbf{p}^r_{-jt}) \right] \end{aligned}$$

Theory

- Problem:
    - Deriving this lower bound implies to compute counterfactual prices and market shares, and thus needs manufacturer's belief about the retailer's reaction if wholesale price changes.
    - Given existence of retailer adjustment costs, it is possible that retail price will not adjust in response to wholesale price change.
- Conjecture:
    - As observed in the data, assume that whenever the wholesaler adjusts, retailer adjusts too and sets a counterfactual retail price p_{jt}^{rc} determined based on FOC of the retailer.

Empirical results

- After demand estimation, compute marginal costs using supply model

Mean prices, markups, and costs for selected foreign brands

	Bass	Becks	Corona	Heineken	All imports
Retailer					
Price	6.36 (0.64)	5.27 (0.48)	5.10 (0.62)	5.66 (0.70)	5.52 (0.90)
Markup	0.41 (0.005)**	0.41 (0.003)**	0.40 (0.004)**	0.40 (0.004)**	0.41 (0.003)**
Non-traded costs					
Backed out	0.26 (0.070)**	0.41 (0.045)**	0.39 (0.091)**	0.47 (0.070)**	0.42 (0.038)**
Fitted	0.24 (0.071)**	0.38 (0.047)**	0.36 (0.091)**	0.45 (0.070)**	0.38 (0.041)**
Manufacturer					
Price	5.76 (0.24)	4.44 (0.16)	4.27 (0.43)	4.99 (0.28)	4.69 (0.77)
Markup	0.51 (0.005)**	0.40 (0.003)**	0.40 (0.003)**	0.53 (0.006)**	0.44 (0.003)**
Total costs					
Backed out	5.28 (0.04)**	4.04 (0.01)**	3.83 (0.07)**	4.59 (0.07)**	4.41 (0.05)**
Fitted	5.21 (0.03)**	4.04 (0.01)**	3.82 (0.05)**	4.48 (0.07)**	4.25 (0.02)**

Empirical results

- Estimation of non-traded retail cost function to get estimates of non-traded costs for periods of non adjustment

$$ntc^r_{jtz} = c_j + \gamma_z d_z + \gamma_w w^d_t + \eta_{jtz}$$

- Derivation of bounds for the retailer price adjustment costs

Bounds for the retailer's adjustment costs for selected foreign brands

	Mean cost		Share of brand's revenue	
Brand	Upper bounds ($)	Lower bounds ($)	Upper bounds (%)	Lower bounds (%)
Bass	$103·32 (58·45)*	$6·16 (2·40)**	3·05 (2·01)	0·18 (0·07)**
Beck's	$456·48 (273·89)*	$143·06 (9·21)**	4·25 (1·30)**	1·33 (0·06)**
Corona	$131·36 (70·95)*	$61·00 (31·99)*	1·09 (0·54)**	0·51 (0·20)*
Heineken	$621·12 (249·73)**	$142·28 (30·00)**	3·50 (1·77)**	0·80 (0·16)**
All	$249·26 (78·96)**	$45·36 (7·97)**	3·13 (1·08)**	0·57 (0·12)**

Empirical results

- Computation of manufacturer marginal costs c_{jt}^{w}
- Estimation of manufacturer marginal cost function when wholesale price don't adjust

$$c_{jt}^{w} = \exp(\theta_j + \omega_{jt})(w_t^d)^{\theta_{dw}}(e_{jt}w_t^f)^{F_j\theta_{fw}}(p_{bjt})^{D_j\theta_{dp}}(e_{jt}p_{bjt})^{F_j\theta_{fp}}$$

  where w_t^d and w_t^f denote local domestic and foreign wages, e_{jt} the bilateral exchange rate between producer country and the US, p_{bjt} the price of barley in the country of production of brand j, F_j a dummy equal 1 if product is produced by foreign supplier, D_j a dummy equal 1 if product produced by domestic supplier

- Function homogeneous of degree 1 in factor prices: $\theta_{dw} + F_j\theta_{fw} + D_j\theta_{dp} + F_j\theta_{fp} = 1$

Empirical results

- Derivation of bounds for the wholesale price adjustment costs A_{jt}^{w}

Bounds for manufacturers' adjustment costs for selected foreign brands

	Mean cost		Share of brand's revenue	
Brand	Upper bounds ($)	Lower bounds ($)	Upper bounds (%)	Lower boun (%)
Bass	$36·76 (18·54)*	$12·86 (0·35)**	0·80 (0·39)*	0·28 (0·02)**
Beck's	$310·44 (121·45)**	$109·14 (21·25)**	2·43 (0·85)**	0·85 (0·15)**
Corona	$213·48 (35·55)**	$181·00 (7·61)**	1·41 (0·23)**	1·20 (0·01)**
Heineken	$122·60 (73·02)**	$112·38 (4·83)**	0·22 (0·10)**	0·20 (0·02)**
All	$178·60 (50·31)**	$85·32 (3·03)**	1·28 (0·34)**	0·61 (0·02)**

Discrete/Continuous Demand

- Consumer choices are Discrete, but sometimes with a lot of modalities
- Discrete choice often criticized because observed consumers often buy several units (cans of soft drinks, ..)
- Several solutions or ways to rationalize data

Discrete/Continuous Demand

- Rationalize multiple choices assuming they are just aggregation over several choice instances
  - For example, a consumer shopping in a store for a week. Assuming each day is a choice decision means consumer who bought 5 cans of soft drinks decided to choose the outside option on two choice occasions.
  - Unappealing, because assumes choices across days are independent.
- Continuous Demand models for homogenous goods (AIDS, ..)
  - Cannot explain discrete dimension
- Discrete/continuous demand model
- Dynamic demand model with stockpiling: quantity choice is not consumption

Discrete/Continuous Demand

- Hendel (1999) studies a multi-discrete choice situation. Observes firms simultaneously buying several brands of computers and several units of each brand.
    - Non-discreteness is in two dimensions:
    - Choice of several brands as an aggregation over several tasks.
    - For each task there is an optimal brand, but observed purchases are aggregation over several tasks.
    - The purchase of several units is explained by a decreasing marginal utility from quantity.

Discrete/Continuous Demand

- Dubin and McFadden (1984) "An econometric Analysis of residential Electric Appliance Holdings and Consumption", Econometrica
- Hanemann (1984) "Discrete/Continuous Models of Consumer Demand", Econometrica

- Objective: Formulating econometric model of discrete/continuous consumer choices in which discrete and continuous choices coming both from the same underlying (random) utility maximization decision

Discrete/Continuous Demand

- Assume preferences $u(x, q, b, z, \epsilon)$ over vector of goods $x_1, .., x_J$, numeraire good q, with goods' attributes in vector b, consumer characteristics z and unobserved taste shocks ϵ
- Assume $u(.)$ is such that $\frac{\partial u}{\partial b_j} = 0$ if $x_j = 0$, such that attributes of good j don't matter if good j not consumed
- Maximize direct utility function $u(x, q, b, z, \epsilon)$ subject to budget constraint:

$$\sum_{j=1}^{J} p_j x_j + q = y$$

and exclusivity constraints

$$x_j x_{j'} = 0 \qquad \forall j \neq j'$$

and positivity constraints on x and q, where x_j is the quantity demanded of good j by the consumer with unit price p_j and q is the outside good (other purchases).

Random Utility Demand Model

- Maximization yields conditional demand functions $\overline{x}_j(p_j, y, b_j, z, \epsilon)$ and $\overline{q}(p_j, y, b_j, z, \epsilon) = y - p_j\overline{x}_j(p_j, y, b_j, z, \epsilon)$
- Conditional indirect utility functions

$$\overline{v}_j(p_j, y, b_j, z, \epsilon) = u(0, .., \overline{x}_j(p_j, y, b_j, z, \epsilon), .., 0, \overline{q}(p_j, y, b_j, z, \epsilon), b, z, \epsilon)$$

- Choice of alternative made from a choice set J, according to:

$$\overline{v}_j(p_j, y, b_j, z, \epsilon) \geq \overline{v}_k(p_j, y, b_j, z, \epsilon) \qquad \forall k \in J$$

- Random for econometrician because unobservable ϵ.
- Specification of the conditional indirect utility function: Dubin and McFadden (1984) Hanemann (1984), Smith (2004), Dubois and Jodar (2012).

Random Utility Demand Model

- Dubois and Jodar (2012) characterize indirect utility functions leading to conditional expenditure function additively separable between income and prices.
- From Roy's identity, $\overline{v}_j$ must satisfy PDE

$$\overline{x}_j(p_j, y, b_j, z, \epsilon) = -\frac{\frac{\partial \overline{v}_j(p_j, y, b_j, z, \epsilon)}{\partial p_j}}{\frac{\partial \overline{v}_j(p_j, y, b_j, z, \epsilon)}{\partial y}}$$

- Several possible specifications

Random Utility Demand Model

- If we impose a linear additive form for conditional expenditure $\overline{e}_j(p_j, y, b_j, z, \epsilon) = p_j \overline{x}_j(p_j, y, b_j, z, \epsilon)$ such as

$$\overline{e}_j(p_j, y, b_j, z, \epsilon) = h(\ln p_j) + \beta_2 y + \omega_j$$

where ω_j may depend on b_j, z, ϵ, and $h(.)$ known increasing

- From Roy's identity, $\overline{v}_j$ must satisfy PDE

$$-p_j \frac{\partial \overline{v}_j}{\partial p_j} / \frac{\partial \overline{v}_j}{\partial y} = h(\ln p_j) + \beta_2 y + \omega_j$$

Random Utility Demand Model

- Thus $\overline{v}_j(p_j, y, b_j, z, \epsilon)$ must be

$$\overline{v}_j = \Phi\left(\left[H(\ln p_j) + \beta_2 y + \psi_j\right] \exp(\phi_2 - \beta_2 \ln p_j), \omega_j\right)$$

with

$$H(z) = \beta_2 \int_z^{+\infty} h(x) \exp \beta_2 (z - x)\, dx$$

and where $\Phi(., \omega_j)$ is increasing in first argument.

- Can show by integration by parts that $H(.)$ is linear if $h(.)$ is affine

$$h(z) = \beta_1 (z - 1/\beta_2) \rightarrow H(z) = \beta_1 z$$

Discrete/Continuous Demand Specification

- Functional form choice (Dubin McFadden, 1984, Smith, 2004, Dubois and Jodar,2012)

$$\overline{v}_{ij} = \left[\beta_2^i y_i + \beta_1^i \ln p_j + \psi_{1j}^i\right] \exp\left[(\phi_{2j}^i - \beta_2^i \ln p_j)\zeta_i\right] + \psi_{2j}^i + \epsilon_{ij}$$

  where ζ_i is an unobservable preference shock and ϵ_{ij} is an unobserved additive preference shock.
- It implies the conditional expenditure function

$$\overline{e}_{ij}(p_j, y_i, b_{ij}, z_i, \epsilon_{ij}) = \left(\beta_1^i \ln p_j + \beta_2^i y_i + \psi_{1j}^i\right)\zeta_i - \frac{\beta_1^i}{\beta_2^i}$$

- ψ_{1j}^i and ψ_{2j}^i are good j specific tastes quality indexes. ψ_{1j}^i can contain observed characteristics z_i.

Discrete/Continuous Demand Specification

- Assuming $(\epsilon_{ij}, \zeta_i) \perp (p_j, y_i, b_j, z_i)$, distribution of $\overline{v}_{ij}$ conditional on (p_j, y_i, b_{ij}, z_i) is non parametrically identified using micro data, but not necessarily that of ζ_i and ϵ_{ij} (Berry and Haile, 2010).
- Assuming that ϵ_{ij} are i.i.d. Type-1 Extreme Value, conditional probability for i to choose j is:

$$s_{ij}(\zeta_i) = \frac{\exp\left\{\left[\beta_2^i y_i + \beta_1^i \ln p_j + \psi_{1j}^i\right] \exp\left[(\phi_{2j}^i \text{-} \beta_2^i \ln p_j)\zeta_i\right] + \psi_{2j}^i\right\}}{\sum_k \exp\left\{\left[\beta_2^i y_i + \beta_1^i \ln p_k + \psi_{1k}^i\right] \exp\left[(\phi_{2k}^i \text{-} \beta_2^i \ln p_k)\zeta_i\right] + \psi_{2k}^i\right\}}$$

- The preference variation multiplicative term $\exp \zeta_i$ gives a random coefficient structure to the random utility.
- Specifying a distribution of coefficients, we can integrate $s_{ij}(\zeta_i)$ as in random coefficient logit model (Berry 1994, Berry et al., 1994).

Discrete/Continuous Demand Specification

- β_1^i determines price elasticity of conditional demand $-1+\beta_1^i \frac{\zeta_i}{x_j p_j}$
- Unconditional demand of good j is

$$s_{ij}\left(\zeta_i\right)\overline{x}_{ij}(p_j, y_i, b_j, z_i, \epsilon_{ij})$$

  depends on all prices and characteristics
- Assumptions: (ζ_i, ϵ_{ij}) i.i.d. ζ_i assumed log-normal $LN(0,\lambda)$ and ϵ_{ij} extreme value
- Then can use Maximum Likelihood

Discrete/Continuous Demand Estimation

- Unconditional choice probability or market share of good j is

$$\pi_j = \int_\zeta s_{ij}\left(\zeta_i\right) dF\left(\zeta_i\right)$$

- The distribution of the unobserved individual characteristics determine a density $f_{\overline{e}_{ij}}\left(\overline{e}_{ij} \mid j\right)$ for this expenditure level:

$$f_{\overline{e}_{ij}}\left(\overline{e}_{ij} \mid j\right) = \frac{1}{\left[\overline{e}_{ij} + \frac{\beta_1^i}{\beta_2^i}\right]\sqrt{2\pi\lambda}} \exp \frac{-1}{2\lambda}\left[\ln\left(\frac{\overline{e}_{ij} + \left(\beta_1^i / \beta_2^i\right)}{\beta_2^i y_i + \beta_1^i \ln p_j + \psi_{1j}^i}\right)\right]^2$$

- So the Likelihood function is:

$$\ln L = \sum_{i,j} d_{ij}\left\{\ln \pi_{ij} + \ln\left[f_{\overline{e}_{ij}}\left(\overline{e}_{ij} \mid j\right)\right]\right\}$$

Discrete/Continuous Demand

- Application (Dubois and Jodar, 2012)
- Consumer decides:
  - In which retailer to buy a bundle of products (assume one-stop shopping model)
  - How much to spend in the bundle
- Dubois and Jodar (2012) analyze selection of "first-choice" retailer (as in Smith, 2004), i.e. the retailer in which most of the expenditure is made, and the conditional expenditure function

Discrete/Continuous Demand

- Firms profit maximization

$$\Pi_h = \sum\nolimits_{j \in J_h} (p_j - c_j)\, q_j$$

where $q_j = \int s_{ij} x_{ij} dF(\zeta_i)$ is the total demand at retail store j.

- First order conditions, take into account effects of prices on conditional demand and on store choice

$$\begin{aligned} 0 &= \int s_{ij} x_{ij} dF(\zeta_i) \\ &\quad + (p_j - c_j) \int s_{ij} \frac{\partial x_{ij}}{\partial p_j} dF(\zeta_i) + \sum_{k \in J_h} (p_k - c_k) \int x_{ik} \frac{\partial s_{ik}}{\partial p_j} dF(\zeta_i) \end{aligned}$$

because $\frac{\partial x_{ik}}{\partial p_j} = 0$ for all $k \neq j$.

Discrete/Continuous Demand

- Model is then used to perform counterfactual simulations
- Evaluate the impact of shock in transportation cost (distance)
- Affects degree of competition between retailers because affects store choice decision and retailers equilibrium price

Demand Model in Product and Characteristic Space

- "Do Prices and Attributes Explain International Differences in Food Purchases", Dubois, Griffith, Nevo, American Economic Review 2014

Motivation

- Food purchases differ across countries, time and demographics
- Seem to be correlated with obesity and obesity-related health outcomes
- The differences are due to many factors
    - economic environment: prices, product attributes
    - other factors: culture, eating habits, preferences
- How far can prices and attributes go in explaining observed differences?

General Strategy

- Document cross-country differences
- Propose and estimate a model of demand for food (and nutrients)
  - build on Gorman (1956); nests product and characteristics models
  - estimate the model for each country separately
- Conduct counterfactuals
  - purchases (and implied nutrients) if faced with prices and attributes from other countries
- Data : France, UK and the US we have:
  - a panel of household purchases gathered using home scanners
  - product level nutritional information

Why Do We Care?

- Differences in nutritional intake mirrored in a number of health outcomes
- For example, obesity rates
  - countries: France: 14.5%, UK: 23.6% ,US: 30.0%
- A key cause of obesity is caloric intake
- Poor nutrition not just about obesity
- Just to be clear
  - other factors are at play: not just food at home
  - we will not directly look at obesity or other health outcomes

Key Findings

- Economic environment is quite important
  - American faced with French prices/attributes: prices diff explain US-FR cals diff
  - American faced with British prices/attributes: prices/attributes matter, but go the "wrong" way
- Prices/attributes are not everything
  - cannot explain composition of consumption, even when agg match
  - cannot explain US-UK differences
- Ranking "health" of the environment: $FR \succ US \succ UK$
  - France has expensive high nutrient products
  - UK has cheap low nutrient products

Related Literature

- Health Economics: diff in obesity across time and markets
  - e.g., Cutler et al. (2003), Philipson and Posner (2003)
  - we quantify the effects
- Development: Deaton (1997) Atkin (2013)
- Nutritional literature
  - e.g., Drewnowski and coauthors show that energy dense food negatively correlated with price
  - Our contribution
    1. use more detailed data
    2. account for several nutrients
    3. estimate a demand curve
    4. conduct counterfactuals

Data

- Data from US, UK and France collected using home scanning devices
  - from market research firms (TNS and Nielsen)
  - participating HH record all food purchased
  - exact date and location of purchase
  - UPC level quantity and price
  - In total hundreds of millions of transactions
- Detailed demographic information
- Nutritional information:
  - information contained on the nutritional label on the back of the package, very detailed

Data Matching: across countries

- The products, and even categories, differ widely across countries;
- We therefore classified the products into 52 categories used in the past by the USDA
- We further collapsed these into 9 broad product categories, which we focus on today
  - Fruits, Vegetables, Grains, Dairy, Meat, Fat, Sugar, Drinks, Prepared Foods
- We focus on 2005-2006

Data Caveats

- Many advantages of these data
    - detailed, comparable across countries
    - panel: observe household over time
    - exact prices and products
- But also some potential concerns about the data:
    - recording error
    - sample selection
    - consumption outside the home
- Try to make comparable as possible, performed several data quality tests when possible

Descriptive Statistics

Table 1 : Demographics

	France	UK	US
# of households	11,677	12,698	8,484
Household size	2.7	2.6	2.4
# of kids	0.7	0.6	0.5
Adult equivalent	2.2	2.1	2.0

Equivalence Scale

- Data are at the household level; household composition varies
- We therefore generate an equivalence scale
  - For each household we compute total caloric needs based on age and gender
    - girl age 4-6 requires 1545 calories per day;
    - boy age 11-14 requires 2220;
  - divide by 2500 to get "adult equivalent"
- There are alternatives that could be explored

Aggregate Purchases

Table 2 : Mean Purchases Across Countries

	FR	UK	US
calories	1776.6	1928.9	2102.7
from carbohydrates	667.4 (38%)	890.5 (47%)	1019.3 (49%)
from protein	287.9 (16%)	293.3 (16%)	264.9 (13%)
from fats	821.0 (46%)	694.5 (37%)	781.6 (37%)
carbohydrates (g)	178.0	237.5	271.8
protein (g)	72.0	73.3	66.2
fats (g)	91.2	77.2	86.8

Average per person per day using an adult equivalent scale

Expenditure and quantity

Table 3: Expenditure and Quantity by Category

	Exp Shares (%)			Quantity			Calorie Share (%)		
Category	FR	UK	US	FR	UK	US	FR	UK	US
Fruits	6.6	9.3	8.1	14.6	14.0	17.2	4.5	4.5	5.3
Vegetables	9.7	10.4	7.9	18.2	20.2	14.0	5.3	6.0	3.0
Grains	6.0	8.4	7.8	6.7	13.4	8.8	14.3	19.8	14.3
Dairy	16.7	12.7	9.5	25.7	27.9	20.7	17.2	12.8	9.3
Meats	31.0	18.3	19.0	14.2	11.1	14.7	16.6	13.2	16.1
Oils	3.3	2.0	1.9	3.1	2.1	2.2	13.1	6.8	6.6
Sweeteners	1.4	1.1	1.4	2.4	2.4	2.6	5.1	4.9	4.4
Drinks	5.9	5.8	10.1	43.4	17.4	50.0	3.5	2.0	5.9
Prepared	21.2	32.7	36.1	16.4	26.2	30.0	22.8	31.2	38.0

Average per person per quarter using an adult equivalent scale, conditional on strictly positive expenditure

Prices

Table 4: Mean Prices by Category

	FR	UK	US
Fruits	2.09	3.21	2.12
Vegetables	2.53	2.32	2.64
Grain	**3.89**	**2.63**	**3.73**
Dairy	**3.26**	**2.22**	**2.48**
Meats	**10.33**	**7.29**	**5.88**
Oils	5.19	3.97	4.47
Sweeteners	2.79	2.38	4.61
Drinks	0.89	2.50	1.56
Prepared	**6.04**	**5.43**	**5.13**

Notes: units are US$ per 1 kilogram

Descriptive Statistics

Table 5: Nutritional Content by Category

calories from:	carbohydrates			protein			fat		
	FR	UK	US	FR	UK	US	FR	UK	US
Fruits	57	68	70	3	5	2	8	7	1
Vegetables	39	38	49	20	22	13	76	85	7
Grain	211	129	227	34	22	38	96	20	36
Dairy	18	22	29	71	57	48	188	166	130
Meats	5	21	30	76	72	66	120	129	206
Oils	2	7	6	11	3	2	678	602	671
Sweeteners	305	307	345	3	4	0	0	1	0
Drinks	27	34	69	1	4	2	1	4	5
Prepared	126	95	194	24	23	22	127	88	117

Average calories from nutrients (carbohydrates, proteins, fats) per 100 grams of food.

Model Overview

- Key challenge: how do we take advantage of the richness of the data?
- Option 1: Estimate demand the "usual" way.
  - is the disaggregated choice relevant for the big picture?
  - can we generalize?
  - how do compare brands across countries?
- Option 2: Use more aggregate product definition
  - how to use nutrient information?
  - how to deal with product differences across countries?
- We follow the second approach and offer a demand system that combines models in product and characteristics space

The Consumer's Problem

- The consumer chooses from N products;
- Product n is characterized by C characteristics $\{a_{n1}, ..., a_{nC}\}$.
- The utility of consumer i with demographics η_i is $U(x_i, \mathbf{z}_i, \mathbf{y}_i; \eta_i)$
  - x_i is the numeraire; $\mathbf{z}_i$ characteristics, $\mathbf{y}_i$ quantities consumed
- Define the $N \times C$ matrix $\mathbf{A} \equiv \{a_{nc}\}_{n=1,..N,c=1,..,C}$
- $U(x, \mathbf{y}, \mathbf{z})$:
  - relies on both characteristics and "flexible" functional forms to guide substitution patterns
  - breaks assumptions of weak separability between food groups
  - control for difference across countries in products

The Consumer's Problem

The consumer's problem:

$$\max_{x_i, \mathbf{y}_i} \quad U(x_i, \mathbf{z}_i, \mathbf{y}_i; \eta_i)$$

$$s.t. \quad \sum_{n=1}^{N} y_{in} p_n + p_0 x_i \leq I_i \quad ; \qquad \mathbf{z}_i = \mathbf{A}' \mathbf{y}_i; \quad x_i, \; y_{in} \geq 0,$$

Dropping i subscripts, then the FOC if $y_n > 0$

$$\sum_{c=1}^{C} a_{nc} \frac{\partial U}{\partial z_c} - \frac{\partial U}{\partial x} \frac{p_n}{p_0} + \frac{\partial U}{\partial y_n} = 0 \quad ,$$

Discussion

- Private case I: $U(x, \mathbf{z})$ – the characteristics demand model
    - With linear technology at most C products consumed
    - Nests: discrete choice, hedonics
- Private case II: $U(x, \mathbf{y})$ – product level demand
    - can be viewed as a characteristics model where:
        - each product has a *unique* characteristic
        - this characteristic is fixed over time/market
- The model $U(x, \mathbf{y}, \mathbf{z})$ can:
    - rely on both characteristics and "flexible" functional forms to guide substitution patterns
    - will break assumptions of weak separability
    - control for difference across countries in products

The Role of Hedonic Prices

- In general

$$\frac{\partial U/\partial y_n}{\partial U/\partial x} = \frac{p_n}{p_0} - \sum_{c=1}^{C} a_{nc} \frac{\partial U/\partial z_c}{\partial U/\partial x}.$$

- If $\frac{\partial U/\partial z_c}{\partial U/\partial x}$ are constants, then this defines (Marshallian) demand, by substituting the "hedonic price" for the price.
- In our model demand depends on the hedonic prices of each good instead of prices
- If two goods have the same price, but one has more of a characteristic that the consumer values positively, they will adjust downwards the hedonic price - the good is more valuable to them

Functional Forms

- Assume J categories, each with K_j products
- Functional form (for now):

$$U(x_i, \mathbf{z_i}, \mathbf{y_i}; \eta_i) = \prod_{j=1}^{J} \left(\sum_{k=1}^{K_j} f_{ikj}\left(y_{ikj}\right) \right)^{\mu_{ij}} \prod_{c=1}^{C} h_{ic}\left(z_{ic}\right) \exp\left(\gamma_i x_i\right)$$

where $z_{ic} = \sum_{k,j} a_{kj,c} y_{ikj}$, $f_{ikj}(y_{ikj}) = \lambda_{ikj} y_{ikj}^{\theta_{ij}}$ and $h_{ic}\left(z_{ic}\right) = \exp(\beta_c z_{ic})$

- Basically, Cobb-Douglas across food groups, and CES aggregator within a group
- Very rich heterogeneity; limited substitution and income effects

Functional Forms

- Maximizing utility subject to budget constraint yields first order conditions:

$$\mu_{ij}\frac{f'_{ikj}\left(y_{ikj}\right)y_{ikj}}{\sum_l f_{ijl}\left(y_{ilj}\right)}+\sum_c a_{kj,c}y_{ikj}\frac{h'_{ic}\left(z_{ic}\right)}{h_{ic}\left(z_{ic}\right)}=\gamma_i\frac{p_{kj}}{p_0}y_{ikj}.$$

for each k,j.

- Summing over k for a given j :

$$\mu_{ij}\frac{\sum_k f'_{ikj}\left(y_{ikj}\right)y_{ikj}}{\sum_k f_{ikj}\left(y_{ikj}\right)}+\sum_c\frac{h'_{ic}\left(z_{ic}\right)}{h_{ic}\left(z_{ic}\right)}\sum_k a_{kj,c}y_{ikj}=\gamma_i\sum_k\frac{p_{kj}}{p_0}y_{ikj}.$$

- Using $f_{ikj}(y_{ikj})=\lambda_{ikj}y_{ikj}^{\theta_{ij}}$ and $h_{ic}\left(z_{ic}\right)=\exp(\beta_c z_{ic})$, we obtain

$$\sum_k p_{kj}y_{ikj}=p_0\frac{\mu_{ij}\theta_{ij}}{\gamma_i}+\sum_c p_0\frac{\beta_c}{\gamma_i}\sum_k a_{kjc}y_{ikj}.$$

Defining Products

- In principle, products could be defined very narrowly
- However this creates several problems
  - Is this a good model for the choice between narrowly defined products?
  - Different characteristics can be at play at different levels
  - We need to make the estimates transferable across countries
- Therefore, focus on J "categories" (the 9 we showed above) each with K_j mutually exclusive products

Estimating Equation

- Assume one characteristic unobserved. Let

$$p_0 \frac{\mu_{ij}\theta_{ij}}{\gamma_i} + p_0 \frac{\beta_1}{\gamma_i} \sum_k a_{kj,1} \times y_{ikjt} = \delta_{ij} + \xi_{jt} + \varepsilon_{ijt}$$

- Normalize, $\gamma_i = 1$ and $p_0 = 1$

$$w_{ijt} = \sum_c \beta_c z_{ijct} + \delta_{ij} + \xi_{jt} + \varepsilon_{ijt}$$

- where
  - $w_{ijt} = \sum_k p_{ikjt} y_{ikjt}$, $z_{ijct} = \sum_k a_{kj,c} y_{ikjt}$
  - δ_{ij} HH-cat FE; ξ_{jt} cat-qtr FE

Identification

- The error term includes:
  - individual preferences for specific categories
  - category specific seasonal effects
  - promotional activities
  - random noise (to expenditure)
- The independent variable, z_{ijct}, likely correlated with these
  - e.g., positive shock ... consume more ... nutrients increase
- Include HH-category and category-quarter FE
  - absorb the mean; remaining error is deviation
  - error will be uncorrelated with regressor if it does not impact quantity choice (or does so in a way that the total nutrients are constant)

Instruments

- The independent variable, $z_{ijct} = \sum_k a_{kj,c} y_{ikjt}$, likely correlated with error term
  - e.g., positive shock ... consume more ... nutrients increase
- Basic idea: assume nutrients of **available** products are (conditionally) exogenous
- To get variation across HH and time, construct

$$\omega_{ijct} = \frac{1}{\#\mathcal{A}_{ijt}} \sum_{k \in \mathcal{A}_{ijt}} a_{kj,c}$$

where $\mathcal{A}_{ijt}$ is the choice set of products in j for type i during t and assume that

$$E\left(\varepsilon_{ijt} | \omega_{ijct}, \delta_{ij}, \xi_{jt}\right) = 0.$$

where types are by area, quarter, preferred store.
- Requires that
  - nutrients (conditional) mean ind of unobserved characteristics
  - error term correlated with intensive but not extensive margin

Demand Estimates

Table 6: IV Estimates: preferences for nutrients

	FR	UK	US
Carbohydrates	1.483***	1.751***	1.459***
Proteins			
Dairy and Meat	23.62***	18.09***	19.79***
Prepared	16.21***	18.36***	51.84***
Other	2.948	2.660	-1.058
Fats			
Dairy and Meat	4.376***	1.315**	1.184
Prepared	10.94***	10.56***	-2.224***
Other	1.748***	3.707***	1.661***

Product Effects

Table 7: IV Estimates: product effects

	FR	UK	US
Fruits	26.92	38.77	31.19
Vegetables	40.77	41.03	32.93
Grains	16.49	18.41	23.54
Dairy	24.66	25.35	14.87
Meat	73.40	38.25	28.89
Oils	10.71	3.11	5.39
Sweeteners	2.53	0.50	1.77
Drinks	24.33	22.67	37.31
Prepared	54.75	73.81	71.45

Average of the household-category and category-quarter fixed effects

Counterfactuals

- Our goal is to simulate what consumers would buy if faced with prices and attributes from other countries
- Focus on US consumers (average, low income, high calorie)
- Will simulate a change in nominal and "real" prices, attributes and unobserved attributes
- Choice of products within category (i.e., average characteristics).

Simulation

- Define the "average" household

$$\overline{\sigma}_j^H = \bar{y}_j \left[\overline{p}_j^H - \sum_c \widehat{\beta}_c^H \overline{a}_{jc}^H \right] \qquad \text{for } H \in \{US, FR, UK\}$$

- 5 parameters in simulation

$$\widehat{y}_j^{(V_1 \dots V_5)} = \frac{\overline{\sigma}_j^{V_3}}{\frac{\overline{p}_j^{V_2}}{\hat{\tau}^{V_5}} - \sum_c \widehat{\beta}_c^{V_4} \overline{a}_{jc}^{V_1}} \qquad V_i \in \{FR, UK, US\}$$

$$\hat{\tau}^V = p_0^V / p_0^{US}$$

Simulated Scenarios

- We consider a US consumer facing UK/FR prices/attributes
- Five counterfactual scenarios

Scenario A: hold fixed US quantities; change attributes;
Scenario B: change prices; simulate choice;
Scenario C: change prices and attributes; simulate choice;
Scenario D: change "real" prices and attributes;
Scenario E: change "real" prices, attributes, and category FE

Table 8: US Consumers Facing French Prices and Attributes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
scenario:		A	B	C	D	E	
attributes	US	FR	US	FR	FR	FR	FR
prices	US	US	FR	FR	FR	FR	FR
product effects	US	US	US	US	US	FR	FR
nutrient pref	US	US	US	US	US	US	FR
price adjustment	1	1	1	1	1.079	1.079	1
Calories	2212.3	2158.3	1884.1	1839.9	2088.5	1946.6	1873.4
Carb (cal)	1092.6	903.0	1166.6	949.4	1070.5	766.6	709.6
	49.4	41.8	61.9	51.6	51.3	39.4	37.9
Prot (cal)	279.40	326.66	171.57	213.18	243.26	287.91	299.90
	12.6	15.1	9.1	11.6	11.6	14.8	16.0
Fat (cal)	840.3	928.7	545.9	677.3	774.8	892.1	863.9
	38.0	43.0	29.0	36.8	37.1	45.8	46.1

Table 9: US Consumers Facing UK Prices and Attributes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
scenario:		A	B	C	D	E	
attributes	US	UK	US	UK	UK	UK	UK
prices	US	US	UK	UK	UK	UK	UK
product effects	US	US	US	US	US	UK	UK
nutrient pref	US	US	US	US	US	US	UK
price adjustment	1	1	1	1	1.089	1.089	1
Calories	2212.3	2015.3	2336.4	2157.3	2567.6	2372.5	1972.8
Carb (cal)	1092.6	936.1	1269.9	1095.9	1293.3	1066.5	926.1
	49.4	46.4	54.4	50.8	50.4	45.0	46.9
Prot (cal)	279.40	313.11	270.36	299.69	359.30	387.66	306.79
	12.6	15.5	11.6	13.9	14.0	16.3	15.6
Fat (cal)	840.3	766.1	796.1	761.8	915.0	918.4	739.9
	38.0	38.0	34.1	35.3	35.6	38.7	37.5

Table 10: Low Income US Consumers Facing French Prices and Attributes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
scenario:		A	B	C	D	E	
attributes	US	FR	US	FR	FR	FR	FR
prices	US	US	FR	FR	FR	FR	FR
product effects	US	US	US	US	US	FR	FR
nutrient pref	US	US	US	US	US	US	FR
price adjustment	1	1	1	1	1.079	1.079	1
Calories	2128.9	2105.5	1805.9	1775.4	2057.1	1639.0	1594.1
Carb (%)	50.5	42.7	64.9	53.2	52.8	41.3	39.3
Prot (%)	12.0	14.4	8.1	10.7	10.8	13.4	15.0
Fat (%)	37.5	42.9	27.0	36.1	36.4	45.3	45.7

Table 11: Low Income US Consumers Facing UK Prices and Attributes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
scenario		A	B	C	D	E	
attributes	US	UK	US	UK	UK	UK	UK
prices	US	US	UK	UK	UK	UK	UK
product effects	US	US	US	US	US	UK	UK
nutrient pref	US	US	US	US	US	US	UK
price adjustment	1	1	1	1	1.089	1.089	1
Calories	2128.9	1891.4	2130.3	1954.8	2407.8	2304.4	1841.3
Carb (%)	50.5	46.7	57.8	53.2	53.1	45.1	47.5
Prot (%)	12.0	15.0	10.3	12.8	12.8	16.0	15.0
Fat (%)	37.5	38.3	32.0	34.0	34.1	38.9	37.5

Table 12: High Calorie US Consumers Facing French Prices and Attributes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
scenario:		A	B	C	D	E	
attributes	US	FR	US	FR	FR	FR	FR
prices	US	US	FR	FR	FR	FR	FR
product effects	US	US	US	US	US	FR	FR
nutrient pref	US	US	US	US	US	US	FR
price adjustment	1	1	1	1	1.079	1.079	1
Calories	3460.0	3239.2	2686.6	2475.2	2805.5	2995.2	2875.0
Carb (%)	49.5	41.0	62.9	51.3	51.0	38.1	36.6
Prot (%)	12.2	15.3	8.4	11.4	11.4	15.0	16.2
Fat (%)	38.4	43.7	28.6	37.3	37.6	46.9	47.2

Table 13: High Calorie US Consumers Facing UK Prices and Attributes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
scenario:		A	B	C	D	E	
attributes	US	UK	US	UK	UK	UK	UK
prices	US	US	UK	UK	UK	UK	UK
product effects	US	US	US	US	US	UK	UK
nutrient pref	US	US	US	US	US	US	UK
price adjustment)	1	1	1	1	1.089	1.089	1
Calories	3460.0	3105.1	3510.9	3119.9	3725.6	3372.6	2810.7
Carb (%)	49.5	46.4	55.1	51.8	51.4	45.3	47.2
Prot (%)	12.2	15.3	11.1	13.4	13.5	16.1	15.3
Fat (%)	38.4	38.3	33.8	34.8	35.1	38.6	37.5

Main Findings

- US/FR difference:
  - prices/attributes explain differences in calories
  - prices seem more important
- US/UK difference:
  - preferences explain the differences in calories
  - attributes seem to push in the "right" direction
  - prices go the "wrong" way
- Source of calories quite different: it is the interaction of preference, prices and attributes that explains the cross country differences
- Can rank "healthiness" of preferences and the environment
  - French environment generally encourages healthier purchasing
  - UK environment (especially prices) generates worse outcomes
    - Note: UK consumers purchase less calories than US consumers due to their preferences and *despite* their environment;

Dynamic Demand

- Goods can be stored and consumption delayed
- Data often allow to observe consumer purchases or store sales but not necessarily consumption
- Promotions, price uncertainty give incentives to do some intertemporal optimization. Larger quantity purchased can be due to increased consumption or to storage or both.
- But most demand models are static
- What does it imply for demand estimation?

Dynamic Demand

- Demand for differentiated products: effect of dynamics depends on reasons generating dynamics:
  - Storable products (food, batteries, ..)
  - Durable Products (cars, PCs, ..)
  - Habit formation
  - Switching costs
  - Learning

Dynamic Demand

- Consider storable products:
  - if storage costs are not too large and current price is low relative to future prices (i.e., the product is on sale), incentive for consumers to store and consume in the future. Pesendorfer (2002), Hendel and Nevo (2006 Rand), Perrone (2014) present evidence that consumers store when prices are low.
  - Hendel and Nevo (2006 EMA) extends the discrete choice static models to allow for storability: find that static model overestimates price elasticity and underestimates the cross price effects.

Dynamic Demand

- Consider durable products:
  - Dynamics arise due to similar trade-offs. Transaction costs in resale market of durable goods (for example, because of adverse selection) implies consumer's decision today of whether or not to buy a durable good (and which one) is costly to change in the future and will impact future utility. Consumer makes a purchase, depending on his current holdings of the good and expectations about future prices and attributes of available products.
  - Impact of durable products on static demand estimation depends if repeat purchase or not.

Dynamic Demand

- If no repeat purchases:
  - Distribution of random coefficients likely to change over time as some consumers purchase and exit the market. If prices fall over time, likely that less price sensitive consumers purchase initially.
  - Forward looking consumers have option value to not purchasing today, reflected in value of outside option.
- If repeat purchases (Gowrisankaran and Rysman, 2009):
  - Consumers do not exit. However, consumers who previously purchased have different value of no purchase since can stay with current product. Problem with static estimation that does not account for different outside option values across consumers and over time.
  - When purchasing, consumers have endogenous duration of holding product and it changes valuation of options. Might find optimal to buy inferior option (lower flow utility) but replace quickly with better/cheaper future option.

Dynamic Demand

- Heterogeneity
  - As in static models, heterogeneity is key to explain data and have flexible demand, but some degree of unobserved heterogeneity needs be sacrificed to deal with dimensionality problem.
- Data
  - As in static models, dynamic model can be estimated using consumer level or market level data. Advantages of consumer level data obvious in dynamic setting: allow see how individual behave over time. But hard to collect for durables purchased infrequently, then need use aggregate data.

Dynamic Demand

Implications of dynamic demand

- If consumers can store, need separate short run and long run response to either a temporary or permanent price change.
- For most applications, long run changes matter.
- If permanent price changes:
    - static estimation yields consistent estimates of long run demand responses if we use only permanent price changes and ignore temporary prices changes. But temporary price change are often most or even all variation in prices.

Dynamic Demand

Implications of dynamic demand

- If temporary price changes, static demand estimates:
  - over-estimate own price effects: demand response to sale is attributed to consumption increase (which in static model equals purchase), and not to increase in storage. Decline in purchases after sale coincides with price increase, and mis-attributed to consumption decline.
  - under-estimate cross price effects: during sale, quantity sold of competing products goes down, but static estimation misses additional effect: decrease in quantity sold in the future. When a competing product was on sale in the past, consumers purchased to consume today and therefore, the relevant "effective" cross price is not the current cross price effect.

Model of Consumer Stockpiling

- Consumer purchases brand j in size $x \in \{1, 2, .., X\} : d_{jxt} \in \{0, 1\}$
- Per-period utility from consumption

$$u_i(c_t, \nu_t) + \alpha_i m_t$$

where c_t is vector of J quantities of each brand, ν_t is vector of J shocks that change the marginal utility of consumption and m_t is utility from outside good.

- Cost of holding vector of inventories by brand i_t is $C_i(i_t)$

Model of Consumer Stockpiling

- Purchase and consumption decisions $\{c, j, x\}$ to maximize

$$\sum_{t=1}^{\infty} \delta^{t-1} E\left[u_i\left(c_t, \nu_t\right) - C_i\left(i_t\right) + a_{jt}\beta_i - \alpha_i p_{jt} + \xi_{jxt} + \varepsilon_{ijxt} | s_1\right]$$

  s.t. $i_t \geq 0$, $c_t \geq 0$, $i_{t+1} = i_t + \sum_x d_{jxt} x_t - c_{jt}$, $\sum_{j,x} d_{jxt} = 1$, where s_t is the information set at t
- No physical depreciation of products
- Decision is made each period with perfect knowledge of current prices
- State variable $s_t = (i_t, p_t, a_t, \nu_t, \varepsilon_t)$: very large dimension
- Forward looking consumer expectations about future price and characteristics of products (could also be changing), future shocks ν_t, ε_t, using information of current and past values of all products attributes
- Without first order Markov assumption, state space even larger

Model of Consumer Stockpiling

- Value function $V_i(s_t)$ can be obtained as the unique solution of a Bellman equation:

$$\begin{aligned} V_i(s_t) &= \max_{\{c,j,x\}} \{ u_i\left(c_t, \nu_t\right) - C_i\left(i_t\right) + a_{jt}\beta_i - \alpha_i p_{jt} + \xi_{jxt} + \varepsilon_{ijxt} \\ &\quad + \delta \int V_i(s_{t+1}) dF_s\left(s_{t+1} | s_t, c, j, x\right) \} \end{aligned}$$

where F_s represents the transition probability of vector of state variables

Model of Consumer Stockpiling

- Given (ν_t, ε_t) i.i.d. over time, reduce dimensionality using integrated value function

$$EV_i(i_t, p_t) \equiv \int V_i(s_t) dF_\varepsilon(\varepsilon_t) dF_\nu(\nu_t)$$

which is unique solution to the integrated Bellman equation

$$\begin{aligned} & EV_i(i_t, p_t) \\ = \; & \max_{\{c,x\}} \int \ln \sum_j \exp \left\{ \begin{array}{l} u_i(c_t, \nu_t) - C_i(i_t) + a_{jt}\beta_i - \alpha_i p_{jt} + \xi_{jxt} \\ + \delta E\left[EV_i(i_{t+1}, p_{t+1}) \,|\, i_t, p_t, c, j, x \right] \end{array} \right\} dF(\nu_t) \end{aligned}$$

Model of Consumer Stockpiling

- Assume brands perfect substitutes in consumption and storage

$$u_i\left(c_t, \nu_t\right) = u_i\left(\sum_j c_{jt}, \sum_j \nu_{jt}\right) \quad C_i\left(i_t\right) = C_i\left(\sum_j i_{jt}\right)$$

- Reduces state space but also modify dynamic problem (modelling tension: brands differentiated at purchase but not consumption)
- Then, can show optimal consumption depends on size but not brand

$$c_k^*\left(s_t; x, k\right) = c_j^*\left(s_t; x, j\right) = c^*\left(s_t; x\right)$$

Model of Consumer Stockpiling

- Then, use inclusive value to reduce state space dimension

$$
\begin{aligned}
& EV_i\left(i_t, p_t\right) \\
= \quad & \max_{\{c,x\}} \int \ln \sum_x \exp \left\{ \begin{array}{c} u_i\left(c_t, \nu_t\right) - C_i\left(i_t\right) + \omega_{ixt} \\ +\delta E\left[EV_i\left(i_{t+1}, p_{t+1}\right) \middle| i_t, p_t, c, x\right] \end{array} \right\} dF\left(\nu_t\right)
\end{aligned}
$$

where

$$
\omega_{ixt} = \ln\left(\sum\nolimits_j \exp\left\{a_{jxt}\beta_i - \alpha_i p_{jxt} + \xi_{jxt}\right\}\right)
$$

- Problem as a choice between sizes, each with utility given by size-specific inclusive value
- Consumer just has to form expectations about the future inclusive value, or in some cases a low number of inclusive values for subsets of products, rather than expectations about the realizations of all attributes of all products.
- State space is still large and includes all prices needed to compute inclusive value

Model of Consumer Stockpiling

- Assumption: ω_{it} contains all the relevant information in s_t for ω_{it+1} conditional on s_t

$$F\left(\omega_{it+1} | s_t\right) = F\left(\omega_{it+1} | \omega_{it}\left(p_t\right)\right)$$

- Can test if other statistics of price matter
- This is consumer specific
- Then, can show expected future value only depends on lower dimensional statistic of full state

$$EV_i\left(i_t, p_t\right) = EV_i\left(i_t, \omega_{it}\left(p_t\right)\right)$$

Model of Consumer Stockpiling

- Estimation with consumer level data:
  - observe purchases over time, harder to get prices of brands not purchased
  - consumer inventory and consumption decisions not observed
- Can apply similar algorithm to Rust (1987)
  - For given parameters, solve dynamic program to obtain purchases and consumption as function of state variables including unobserved shocks
  - Assuming distribution of shocks, derive likelihood of observing consumer's decision conditional on prices and inventory
  - Search parameters values that maximize observed sample likelihood
  - Problem: state variable inventory not observed and high dimension state space

Model of Consumer Stockpiling

- Assume $\xi_{jxt} = \xi_{jx}$ (use fixed effects)
- Likelihood, given demographics D_i: $P\left(\left(j_1, x_1\right), .., \left(j_T, x_T\right)\right)$ is

$$\int \prod_{t=1}^{T} P\left(j_t, x_t | p_t, i_t\left(d_{t-1}, ., d_1, \nu_{t-1}, ., \nu_1, i_1\right), \nu_t, D_i\right)$$
$$dF\left(\nu_1, ., \nu_T\right) dF\left(i_1\right)$$

$$P\left(j_t, x_t | p_t, i_t, \nu_t, D_i\right) = \frac{\exp\left(a_{jxt}\beta_i \text{-} \alpha_i p_{jxt} + \xi_{jx} + M\left(\omega_t, i_t, \nu_t, j, x\right)\right)}{\sum\limits_{k,y} \exp\left(a_{kyt}\beta_i \text{-} \alpha_i p_{kyt} + \xi_{ky} + M\left(\omega_t, i_t, \nu_t, k, y\right)\right)}$$

where

$$\begin{aligned} & M\left(\omega_t, i_t, \nu_t, j, x\right) \\ = \;& \max_c \left[u_i\left(c_t, \nu_t\right) \text{-} C_i\left(i_t\right) \delta E\left[EV_i\left(i_{t+1}, \omega_{t+1}\right) | i_t, \omega_t, c, j, x\right]\right] \end{aligned}$$

Model of Consumer Stockpiling

- Then, can split likelihood using brand indifference w.r.t. consumption and inventory:

$$M\left(\omega_t, i_t, \nu_t, j, x\right) = M\left(\omega_t, i_t, \nu_t, x\right)$$

then

$$\begin{aligned}
& P\left(j, x \middle| p_t, i_t, \nu_t, D_i\right) \\
= \quad & \frac{\exp\left(a_{jxt}\beta_i \text{-} \alpha_i p_{jxt} + \xi_{jx} + M\left(\omega_t, i_t, \nu_t, j, x\right)\right)}{\sum\limits_{k,y} \exp\left(a_{kyt}\beta_i \text{-} \alpha_i p_{kyt} + \xi_{ky} + M\left(\omega_t, i_t, \nu_t, k, y\right)\right)} \\
= \quad & \frac{\exp\left(a_{jxt}\beta_i - \alpha_i p_{jxt} + \xi_{jx}\right)}{\sum_k \exp\left(a_{kxt}\beta_i - \alpha_i p_{kxt} + \xi_{kx}\right)} \frac{\exp\left(\omega_{xt} + M\left(\omega_t, i_t, \nu_t, x\right)\right)}{\sum_y \exp\left(\omega_{yt} + M\left(\omega_t, i_t, \nu_t, y\right)\right)} \\
= \quad & P\left(j \middle| p_t, x_t, D_i\right) P\left(x_t \middle| \omega_t, i_t, D_i\right)
\end{aligned}$$

Model of Consumer Stockpiling

- Useful only if can assume that heterogeneity only as function of observable demographics

$$F(\alpha_i, \beta_i | x_t, p_t, D_i) = F(\alpha_i, \beta_i | p_t, D_i)$$

then

$$\begin{aligned} P(j|p_t, x_t, D_i) &= \int P(j|p_t, x_t, \alpha_i, \beta_i)\, dF(\alpha_i, \beta_i | x_t, p_t, D_i) \\ &= \int P(j|p_t, x_t, \alpha_i, \beta_i)\, dF(\alpha_i, \beta_i | p_t, D_i) \end{aligned}$$

otherwise need solve the dynamic programming to obtain

$$F(\alpha_i, \beta_i | x_t, p_t, D_i)$$

Model of Consumer Stockpiling

Likelihood in three steps (significantly reduces computational cost)

1. Estimate α_i, β_i by maximizing $P(j|p_t, x_t)$: (static) conditional logit using only options of size x_t. Can include many controls (concerns price endogeneity ..)
2. Use estimated parameters to get ω_{xit} and estimate $F(\omega_{it+1}|\omega_{it})$
3. Estimate dynamic parameters (utility from consumption, storage cost and distribution of ν_t) using $P(x_t|\omega_t, i_t)$ which require solving modified dynamic program

Model of Consumer Stockpiling

TABLE IV

FIRST STEP: BRAND CHOICE CONDITIONAL ON SIZE[a]

	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)	(x)
Price	−0.51	−1.06	−0.49	−0.26	−0.27	−0.38	−0.38	−0.57	−1.41	−0.75
	(0.022)	(0.038)	(0.043)	(0.050)	(0.052)	(0.055)	(0.056)	(0.085)	(0.092)	(0.098)
*Suburban dummy				−0.33	−0.30	−0.34	−0.33	−0.25	−0.45	−0.19
				(0.055)	(0.061)	(0.055)	(0.056)	(0.113)	(0.127)	(0.127)
*Nonwhite dummy				−0.34	−0.39	−0.38	−0.33	−0.34	−0.33	−0.26
				(0.075)	(0.083)	(0.076)	(0.076)	(0.152)	(0.166)	(0.168)
Large family				−0.23	−0.13	−0.21	−0.22	−0.46	−0.38	−0.43
				(0.080)	(0.107)	(0.080)	(0.082)	(0.181)	(0.192)	(0.195)
Feature			1.06	1.05	1.08	0.92	0.93	1.08		1.05
			(0.095)	(0.096)	(0.097)	(0.099)	(0.100)	(0.123)		(0.126)
Display			1.19	1.17	1.20	1.14	1.15	1.55		1.52
			(0.069)	(0.070)	(0.071)	(0.071)	(0.072)	(0.093)		(0.093)
Brand dummy variable		✓	✓	✓	✓					
*Demographics					✓					
*Size						✓				
Brand–size dummy variable							✓			
Brand-HH dummy variable								✓		
*Size									✓	✓

[a]Estimates of a conditional logit model. An observation is a purchase instance by a household. Options include only products of the same size as the product actually purchased
Asymptotic standard errors are shown in parentheses.

Model of Consumer Stockpiling

TABLE V

SECOND STEP: ESTIMATES OF THE PRICE PROCESS[a]

	Same Process for All Types				Different Process for Each Type			
	ω_{2t}	ω_{4t}	ω_{2t}	ω_{4t}	ω_{2t}	ω_{4t}	ω_{2t}	ω_{4t}
$\omega_{1,t-1}$	0.003	−0.014	0.005	0.014	−0.023	−0.005	−0.019	0.007
	(0.012)	(0.011)	(0.014)	(0.014)	(0.017)	(0.014)	(0.019)	(0.015)
$\omega_{2,t-1}$	0.413	0.033	0.295	0.025	0.575	−0.003	0.520	0.011
	(0.007)	(0.010)	(0.008)	(0.007)	(0.013)	(0.010)	(0.016)	(0.013)
$\omega_{3,t-1}$	0.003	−0.034	0.041	−0.006	0.027	−0.072	0.051	−0.018
	(0.007)	(0.007)	(0.009)	(0.009)	(0.020)	(0.016)	(0.025)	(0.020)
$\omega_{4,t-1}$	0.029	0.249	0.026	0.236	−0.018	0.336	−0.018	0.274
	(0.008)	(0.008)	(0.008)	(0.017)	(0.020)	(0.016)	(0.021)	(0.017)
$\sum_{\tau=2}^{5} \omega_{1,t-\tau}$			−0.003	−0.012			−0.008	−0.003
			(0.005)	(0.004)			(0.006)	(0.005)
$\sum_{\tau=2}^{5} \omega_{2,t-\tau}$			0.089	0.006			0.073	−0.004
			(0.003)	(0.002)			(0.005)	(0.004)
$\sum_{\tau=2}^{5} \omega_{3,t-\tau}$			−0.008	−0.009			−0.004	−0.016
			(0.003)	(0.003)			(0.008)	(0.006)
$\sum_{\tau=2}^{5} \omega_{4,t-\tau}$			−0.013	0.018			−0.008	0.056
			(0.003)	(0.003)			(0.007)	(0.005)

Model of Consumer Stockpiling

TABLE VI

THIRD STEP: ESTIMATES OF DYNAMIC PARAMETERS[a]

Household Type:	1	2	3	4	5	6
		Urban Market			Suburban Market	
Household Size:	1–2	3–4	5+	1–2	3–4	5+
Cost of inventory						
Linear	9.24	6.49	21.96	4.24	4.13	11.75
	(0.01)	(0.02)	(0.09)	(0.01)	(0.17)	(5.3)
Quadratic	−3.82	1.80	−35.86	−8.20	−6.14	−0.73
	(29.8)	(1.77)	(0.19)	(0.03)	(1.69)	(1.53)
Utility from consumption	1.31	0.75	0.51	0.08	0.92	3.80
	(0.02)	(0.09)	(0.21)	(0.03)	(0.18)	(0.38)
Log likelihood	365.6	926.8	1,530.1	1,037.1	543.6	1,086.1

[a] Asymptotic standard errors are shown in parentheses. Also included are size fixed effects, which are allowed to vary by household type.

Ratio of Short Run to Long Run Elasticities

TABLE VIII

AVERAGE RATIOS OF ELASTICITIES COMPUTED FROM A STATIC MODEL TO LONG-RUN ELASTICITIES COMPUTED FROM THE DYNAMIC MODEL[a]

		64 oz.						128 oz.					
Brand	Size (oz.)	All[b]	Wisk	Surf	Cheer	Tide	Private Label	All[b]	Wisk	Surf	Cheer	Tide	Private Label
All[b]	64	1.03	0.13	0.14	0.12	0.13	0.15	0.14	0.17	0.17	0.18	0.21	0.34
	128	0.17	0.24	0.26	0.20	0.28	0.35	1.23	0.09	0.11	0.09	0.15	0.22
Wisk	64	0.14	1.20	0.13	0.17	0.12	0.13	0.16	0.22	0.14	0.22	0.25	0.20
	128	0.25	0.27	0.23	0.31	0.26	0.28	0.08	1.42	0.08	0.13	0.18	0.11
Surf	64	0.14	0.13	0.93	0.16	0.13	0.14	0.18	0.18	0.12	0.18	0.22	0.28
	128	0.25	0.22	0.18	0.27	0.25	0.18	0.12	0.11	1.20	0.08	0.15	0.14
Cheer	64	0.12	0.17	0.16	0.84	0.09	0.13	0.14	0.24	0.16	0.14	0.22	0.24
	128	0.25	0.26	0.26	0.12	0.23	0.22	0.09	0.12	0.06	0.89	0.15	0.07
Tide	64	0.16	0.17	0.13	0.13	1.26	0.15	0.22	0.28	0.16	0.26	0.22	0.37
	128	0.25	0.31	0.22	0.24	0.22	0.31	0.11	0.16	0.08	0.13	1.44	0.31
Solo	64	0.15	0.12	0.15	0.14	0.12	0.14	0.17	0.15	0.15	0.30	0.30	0.28
	128	0.23	0.20	0.24	0.21	0.21	0.25	0.07	0.07	0.06	0.16	0.17	0.21
Era	64	0.21	0.12	0.13	0.13	0.10	0.19	0.43	0.17	0.15	0.22	0.19	0.35
	128	0.31	0.22	0.24	0.25	0.17	0.38	0.19	0.08	0.09	0.11	0.10	0.22
Private	64	0.19	0.15	0.14	0.17	0.17	1.02	0.32	0.22	0.15	0.26	0.31	0.25
label	128	0.29	0.28	0.34	0.30	0.39	0.29	0.16	0.12	0.13	0.10	0.27	1.29
No purchase		2.12	1.13	1.15	1.40	1.27	2.39	1.80	7.60	2.26	14.11	2.38	10.86

[a]Cell entries i and j, where i indexes row and j indexes column, give the ratio of the (short-run) elasticities computed from a static model divided by the long-run elasticities computed from the dynamic model. The elasticities for both models are the percent change in market share of brand i with a 1 percent change in the price of j. The static model is identical to the model estimated in the first step, except that brands of all sizes are included as well as a no-purchase decision, not just products of the same size as the chosen option. The results from the dynamic model are based on the results presented in Tables IV–VI.

[b]Note that "All" is the name of a detergent produced by Unilever.

Elasticities

- Static own-price elasticities overestimate dynamic ones by roughly 30%
- Part of this difference is driven by the bias in estimates of static model: price coefficient estimated in static model is roughly 15 percent higher than estimated in the first stage of the dynamic model
- Static cross-price elasticities, with the exception of the no-purchase option, are smaller than the long-run elasticities
- The effect on the no-purchase option is expected because the static model fails to account for the effect of inventory. A short-run price increase is most likely to chase away consumers who can wait for a better price, namely those with high inventories. Therefore, the static model will overestimate the substitution to the no-purchase option

Elasticities

- Several effects impact cross-price elasticities to the other brands:
  - coefficients estimated in static model tend to be upward biased
  - additional effect due to difference between long-run and short-run effects: consider a reduction in price of a brand.
    - Static elasticities computed from temporary price reductions. Switchers are households willing to switch, from another brand and have a low enough inventory at the time of the price change.
    - Long-run elasticities capture those households willing to substitute at all relevant levels of inventory, because they represent reactions to a permanent change in the price of that brand
    - Which one dominates depends on relative size of both effects and whether observed price variation was temporary or more permanent.
- For own-price effect and cross-price effects towards no-purchase option: econometric bias and the difference between short- and long-run effects operate in the same direction. Both overstate price responses.

Usage-Based Pricing and Demand for Residential Broadband

- Increased usage of internet has led to congestion.
- Usage based pricing (three-part tariffs) is a proposed way to fix this problem.
- Welfare implications of this pricing is only theoretical.
- Demand estimation of broadband services to evaluate welfare implications of alternatives to address network congestion.
- Detailed (high frequency) data of a North American ISP.

Usage-Based Pricing and Demand for Residential Broadband

- Three part tariffs: monthly fee for a monthly data allowance. Price per Gigabyte if allowance exceeded.
- Marginal price is 0 until allowance is used.
- Forward looking users realize shadow price is not 0.
- Shadow price depends on days left on the billing cycle and fraction of allowance already used.
- Different plans offered: from almost linear prices to very high allowances.
- These features imply the need for a dynamic model of consumers decisions throughout the billing cycle.

Data

- 54.801 subscribers from a North American ISP. Alternative is a slow DSL connection.
- Features of plan offered: monthly fees, maximum download and upload speeds, usage allowance and overage price per GB.
- ISP offers only usage based plans to new subscribers. Some old subscribers have unlimited plans.
- Monthly usage from may 2011 to may 2012.
- 15-minutes intervals from may 10th to June 30th, 2012.
- Consumers can carefully track their usage data

Data

- Rapid growth in usage. Median subscriber's usage more than doubles.
- Cyclical pattern during the day. Peak usage between 10 and 11pm.
- On average unlimited plan users use more internet and pay less per GB.
- Subscribers choose optimal plans:
  - 10% of consumers exceeded their allowance in June 2012.
  - Dominated plans: could have paid less for service no slower.
  - Considering 13 months, only 0.13% of consumers chose dominated plan.

Consumers

- Identification depends on consumers responding to shadow prices.
- Marginal cost for consumers exceeding allowance: discounted overage price times the probability of exceeding the allowance.
- Heavy users: behave as shadow price equals to overage price.
- Consumers not exceeding allowance: behavior should not vary during cycle.
- Consumers in between: usage should vary depending on the day and the cumulative usage so far.
- Estimate regression of usage on allowance used so far and the current day of the billing cycle:
    - Current usage is responsive to past usage (within billing cycle).
    - Discrete change in the shadow price when usage is refreshed (across billing cycle).

Model

- Plans are indexed by k, have speed s_k, allowance $\overline{C}_k$, fixed fee F_k and per-GB overage price p_k.
- Consumers have quasi-linear utilities, where subscriber of type h on plan k utility is:

$$u_h(c_t, y_t; v_t; k) = v_t \left(\frac{c_t^{1-\beta_h}}{1-\beta_h} \right) - c_t \left(\kappa_{1h} + \frac{\kappa_{2h}}{\ln(s_k)} \right) + y_t$$

- The first term is the gross utility from content. v_t is a time-varying unobservable not known by subscriber until period t. For each type h, v_t is i.i.d $LN\left(\mu_h, \sigma_h\right)$.
- The second term is the non-price cost of consuming online content.
- κ_{1h} is opportunity cost of content (wait time), and the ratio of κ_{2h} is preference for speed.
- The specification implies a satiation point.

Model

- A vector of parameters $(\beta_h, \kappa_{1h}, \kappa_{2h}, \mu_h, \sigma_h)$ describes a subscriber of type h
- Conditional on choosing plan k, subscriber's problem is:

$$\max_{c_1,..,c_T} \sum_{t=1}^{T} E\left[u_h\left(c_t, y_t, v_t; k\right)\right]$$

$s.t.$

$$F_k + p_k \max\left(C_T - \overline{C}_k, 0\right) + Y_T \leq I$$

$$C_T = \sum_{t=1}^{T} c_t \quad , \quad Y_T = \sum_{t=1}^{T} y_t$$

- Assume that I is large enough so wealth don't constrain consumption of content.
- Stochastic finite horizon dynamic program.

Model

- Value function and optimal policy function:

$$V_{hkt}\left(C_{t-1}, v_t\right) = \left\{ v_t \left(\frac{c_t^{1-\beta_h}}{1-\beta_h} \right) - c_t \left(\kappa_{1h} + \frac{\kappa_{2h}}{\ln(s_k)} \right) + y_t \right.$$
$$\left. - p_k O_{tk}(c_t) + E\left[V_{hk(t+1)}\left(C_{t-1} + c^*_{hkt}\right) \right] \right\}$$

where overage is $O_{tk}(c_t) = \max(c_t - \overline{C}_{kt})$ and $\overline{C}_{kt} = \max(\overline{C}_k - C_{t-1}, 0)$ and

$$c^*_{hkt}\left(C_{t-1}, v_t\right) = \arg\max V_{hkt}\left(C_{t-1}, v_t\right)$$

- Define the shadow price:

$$\widetilde{p}_k(c_t, C_{t-1}) = \left\{ \begin{array}{cl} p_k & \text{if } O_{tk}(c_t) > 0 \\ \frac{dE\left[V_{hk(t+1)}(C_{t-1}+c_t)\right]}{dc_t} & \text{if } O_{tk}(c_t) = 0 \end{array} \right.$$

Model

- Solve for the consumer's optimal choice in period t as a function of the shadow price and the parameters

$$c^*_{hkt} = \left(\frac{v_t}{\kappa_{1h} + \frac{\kappa_{2h}}{\ln(s_k)} + \widetilde{p}_k(c_t, C_{t-1})} \right)^{\frac{1}{\beta_h}}$$

- Then, the expected value function:

$$E\left[V_{hkt}\left(C_{t-1}\right)\right] = \int_0^{\overline{v}_h} V_{hkt}\left(C_{t-1}, v_t\right) dG_h\left(v_t\right)$$

- And the mean subscriber's usage at each state:

$$E\left[c^*_{hkt}\left(C_{t-1}\right)\right] = \int_0^{\overline{v}_h} c^*_{hkt}\left(C_{t-1}, v_t\right) dG_h\left(v_t\right)$$

- Optimal plan choice:

$$k^*_h = \arg\max_k \left\{ E\left[V_{hk1}\left(0\right) - F_k\right]\right\}$$

Estimation

- First step: solve the model.
  - Solve the model for 18,144 types and store the optimal policy.
  - Solution can be characterized by the expected value functions and policy functions (because subscriber does not now v_t prior to period t).
  - Compute transition probabilities between possible states.
- Second step: empirical and predicted moments.
  - Choose weights minimizing the distance between moments in the data and average predicted moments (what mixture of types results in predicted behavior best matching the data).
  - Two sets of moments: mean usage at each state and mass of subscribers at a particular state.

Results

- Most common type accounts for 43% of total mass, top 10 for 83% and top 30 for 96%.
- Correlation between empirical moments and fitted moments is above 0.99. This means that model replicates average usage and density of subscribers at each state successfully.
- Allowing many types is important.
- Majority of high volume subscribers have highly elastic demand.
- Overall value placed in improving speed is substantial.
- Under linear tariff: average willingness to pay is roughly 280 dollars per month.
- Waiting costs are important: traffic is likely to increase in the future.

Counterfactuals

- Usage based pricing compared to unlimited allowances. Same plans but with no overage prices.
  - Subscriber welfare decreases with UBP.
  - Effect on total welfare is less clear.
  - 20% of consumers choose same plan, 74% choose different plans, 1% purchase only when service is unlimited.
  - Some consumers increase usage under UBP.
  - UBP modestly increases total surplus generated from usage, while transferring some surplus from subscribers to the ISP.

Counterfactuals

Table 7: *Usage-Based Pricing Counterfactual, Usage and Welfare*

	Same Plan Unlimited (1)	Same Plan UBP (2)	Switch Plan Unlimited (3)	Switch Plan UBP (4)	Only Unlimited Unlimited (5)	Only Unlimited UBP (6)
Percent of Types (%)	20.23		73.53		1.31	0
Usage and Surplus						
Speed (Mb/s)	16.23	16.23	10.43	13.05	10	–
Usage (GBs)	77.76	62.57	36.47	34.74	23.74	–
Consumer Surplus ($)	189.25	177.40	153.96	138.13	4.84	–
Revenue ($)	81.07	81.14	43.25	64.48	39.99	–
Total Surplus ($)	270.32	258.54	197.21	202.61	44.83	–
Δ in Total Surplus ($)	-11.78		5.40		44.83	–

	Same Plan ↑ Usage (1)	Same Plan ↓ Usage (2)	Switch Plan ↑ Usage (3)	Switch Plan ↓ Usage (4)	Only Unlimited Unlimited (5)	Only Unlimited UBP (6)
Percent of Types (%)	0	20.23	65.63	7.90	1.31	0
Mean Type						
mean of shocks ($\bar{\mu}_h$)	–	0.91	1.28	0.49	0.24	–
s.d. of shocks ($\bar{\sigma}_h$)	–	0.75	0.82	0.74	0.78	–
opp cost of content ($\bar{\kappa}_{1h}$)	–	4.88	6.67	2.54	2.35	–
pref for speed ($\bar{\kappa}_{2h}$)	–	7.04	5.12	3.56	0.70	–
curvature ($\bar{\beta}_h$)	–	0.44	0.41	0.38	0.41	–

Counterfactuals

- Economic viability of Next Generation Networks.
- Single plan with unlimited usage and one GB/s connection.
- Assume fixed fee of 100 dollars.
- Total surplus is 87 dollars higher per subscriber.
- ISP only realized 22 dollars of additional revenue.
- This implies gap between social and private incentives to invest.
- Socially the investment is recovered in 37 months, but for the ISP in 150 months.

Counterfactuals

Table 8: *Next-Generation Network Counterfactual, Usage and Welfare*

	Both		Single	
	Unlimited	**UBP**	**Unlimited**	**UBP**
	(1)	**(2)**	**(3)**	**(4)**
Percent of Types (%)	84.40		0	9.36
Usage and Surplus				
Speed (Mb/s)	1,024.00	13.97	–	11.74
Usage (GBs)	102.24	41.80	–	31.30
Consumer Surplus ($)	235.97	160.37	–	22.64
Revenue ($)	100.00	70.03	–	50.46
Total Surplus ($)	335.97	230.40	–	73.10
Δ in Total Surplus ($)	-105.57		–	73.10

	Both		Single	
	↑ Usage	**↓ Usage**	**Unlimited**	**UBP**
	(1)	**(2)**	**(3)**	**(4)**
Percent of Types (%)	0	84.40	0	9.36
Mean Type				
mean of shocks ($\bar{\mu}_h$)	–	1.26	–	-0.01
s.d. of shocks ($\bar{\sigma}_h$)	–	0.82	–	0.55
opp cost of content ($\bar{\kappa}_{1h}$)	–	6.17	–	3.82
pref for speed ($\bar{\kappa}_{2h}$)	–	5.69	–	2.86
curvature ($\bar{\beta}_h$)	–	0.40	–	0.56

Conclusion

- Analyze empirically two alternatives to deal with congestion: usage based pricing and high speed next generation networks.
- USB pricing increases total welfare. Higher speed decreases waiting costs. Usage only falls slightly.
- What would happen with plans is important to the welfare implications.
- For the next generation networks, there is a large gap between private and social incentives to invest in those networks.