EMPIRICAL ANALYSIS OF INNOVATION IN OLIGOPOLY INDUSTRIES

Lecture 2: Consumer valuation of product innovations

CEMFI SUMMER SCHOOL – 2018
Victor Aguirregabiria (University of Toronto)

September 4, 2018

1. Introduction

Consumer valuation of product innovations

- Product innovation is ubiquous is most industries, and a key strategy for differentiation.
- During the last decades we have witnessed a large increase in the number of varieties of different products.
- Evaluating consumer value of new products, and of quality improvements in existing products, has received substantial attention in the context of:
  - Improving Cost of Living Indexes (COLI).
  - Costs and benefits of firms' product differentiation.
  - Social value of innovations.

Introduction [2]

- The standard approach is based on:
  - Estimation of a demand system of differentiated products;
  - Constructing consumer indirect utility function (or surplus function)
  with and without the new product.
- Typically, one of the two scenarios (with or without) is a counterfactual.
- In the definition of the counterfactual scenario the researcher needs to **the value of unobservables in the counterfactual scenario.**
  e.g., industry time-trends, unobserved product characteristics, distribution of consumer idiosyncratic product-specific shocks.

Lecture 2: Consumer valuation of product innovations

- 1. Introduction
- 2. Demand for differentiated products
- 3. Valuation of new products
  - 3.1. Trajtenberg (JPE, 1989)
  - 3.2. Petrin (JPE, 2002)
  - 3.3. Valuing new goods with product complementarity: Gentzkow (AER, 2007)

2. Demand for differentiated products

Outline

1. Demand systems in product space
2. Demand systems in characteristics space

Main References

- Nevo (Annual Review of Economics, 2011). Excellent survey.
- Hausman (NBER book, 1996). Application using demand system in product space.
- Berry (RAND, 1994) and Berry, Levinsohn, & Pakes (Econometrica, 1995). Seminal papers on discrete-choice demand-system in characteristics space.
- Nevo (Econometrica, 2001). Application using demand system in product space.
- Berry & Haile (Econometrica, 2014). Results on nonparametric identification.

Demand systems in product space

- Consumer preferences are defined over goods (or varieties) themselves.

- J varieties indexed by $j \in \{1, 2, ..., J\}$. Consumer has a utility function $U(c, q_1, q_2, ..., q_J)$.Consumer problem:

$$\max_{\{c, q_1, q_2, ..., q_J\}} \quad U(c, q_1, q_2, ..., q_J)$$

$$subject\ to: \quad c + p_1\ q_1 + p_2\ q_2 + ... + p_J\ q_J \leq y$$

- Solution to this optimization problem is the **System of Marshallian demand equations**:

$$q_j \quad = \quad f_j\left(p_1, p_2, ..., p_J, y\right)$$

Example 1: Linear Expenditure System

- Consider the Stone-Geary utility function:

$$U = c\ (q_1 - \gamma_1)^{\alpha_1}\ (q_2 - \gamma_2)^{\alpha_2} \ldots\ (q_J - \gamma_J)^{\alpha_J}$$

- For this utility function, the Marshallian demand equations are the so called **Linear Expenditure System**:

$$q_j = \gamma_j + \frac{\alpha_j^*}{p_j} \left(y - \sum_{i=1}^{J} p_i\ \gamma_i \right)$$

  where $\alpha_j^* = \alpha_j / [1 + \sum_{i=1}^{J} \alpha_i]$.

- Convenient for its simplicity but very restrictive. It imposes the restriction that **all the goods are complements in consumption**.

- This is not realistic in most applications, particularly when the goods under study are varieties of a differentiated product.

Example 2: CES Demand System

- Consider the CES utility function:

$$U = c^{\alpha} \left(\sum_{j=1}^{J} q_j^{\sigma} \right)^{1/\sigma}$$

where $\sigma \in [0, 1]$ is a parameter that represents the degree of substitution between the J products.

- This utility function implies the demand system:

$$q_j = \frac{1}{1+\alpha} \left(\frac{P}{p_j} \right)^{\sigma/(1-\sigma)} \left(\frac{y}{p_j} \right)$$

where P is the price index $\left(\sum_{k=1}^{J} p_k^{-\sigma/(1-\sigma)} \right)^{-(1-\sigma)/\sigma}$.

- For any pair of varieties, j and k, we have that:

$$\ln \left(\frac{q_j}{q_k} \right) = \frac{-1}{1-\sigma} \ln \ln \left(\frac{p_j}{p_k} \right)$$

Example 2: CES Demand System [2]

- The CES system imposes strong restrictions on cross-price elasticities. For any three varieties, say j, k, and l:

$$\frac{\partial \ln q_k}{\partial \ln p_j} = \frac{\partial \ln q_l}{\partial \ln p_j}$$

- Suppose that we use this system to study the demand of different varieties of automobiles.
- Suppose that products j and k are "similar" luxury cars, and product l is an basic and expensive variety.
- The CES model implies that a reduction in the price of the luxury car j implies the same proportional increase in the demand of the other luxury car k and the basic car l.
- This is very unrealistic.

Example 3: 'Almost Ideal Demand System'

- Deaton and Muellbauer (1982) proposed the utility function:

$$U = \left[\prod_{j=1}^{J} q_j^{\alpha_j} \right] + \sum_{j=1}^{J} \sum_{k=1}^{J} \delta_{jk} \ q_j \ q_k$$

  that allows for complementarity and substitutability between produces.

- The Marshallian demand equations are:

$$w_j = \alpha_j + \gamma_j \ \ln(y) + \sum_{k=1}^{J} \beta_{jk} \ \ln(p_k)$$

  \- $w_j \equiv p_j q_j / y$ is the expenditure share of product j;
  \- $\{\alpha_j, \beta_{jk}, \gamma_j\}$ are parameters which are known functions of the utility parameters $\{\alpha_j, \delta_{jk}\}$.

- The model implies the symmetry conditions $\beta_{jk} = \beta_{kj}$. Therefore, the number of free parameters is: $2J + \frac{J(J+1)}{2}$, that increases quadratically with the number of products.

Multi-stage Budgeting

- For products with many (> 100) varieties (automobiles, smartphones, cereals, beer, etc) the number of parameters to estimate in the AIDS system can be very large, even larger than the #observations.
- Deaton and Muellbauer (1982) propose using a multi-stage budgeting approach.
- Suppose that the utility function is separable in the utility from G groups of products:

$$U = f\left(v_1\left(\widetilde{\mathbf{q}}_1\right), v_2\left(\widetilde{\mathbf{q}}_2\right), ..., v_G\left(\widetilde{\mathbf{q}}_G\right)\right)$$

  $\widetilde{\mathbf{q}}_g =$ Vector of quantities of varieties in group g;
  $v_g\left(\widetilde{\mathbf{q}}_g\right) =$ Sub-utility from group g;
  $f\left(v_1, ..., v_G\right)$ is an increasing function.

- Assumption: indirect utility functions for each group satisfy Generalized Gorman Polar form. AIDS satisfies this condition.

Multi-stage Budgeting (under AIDS)

- Then, the demand system at the lower stage (**within-group stage**) is:

$$w_{jt} = \alpha_j + \gamma_j \ \ln\left(\frac{y_{gt}}{P_{gt}}\right) + \sum_{k \in \mathcal{J}_g} \beta_{jk} \ \ln(p_{kt})$$

y_{gt} = Expenditure in group g;
P_{gt} = Price index for group g.

- According to the model, this price index depends on the parameters of the model in group g. Non-linear system. Typically applications use "short-cuts": e.g., $\ln P_{gt} = \sum_{j \in \mathcal{J}_g} w_{jt} \ \ln(p_{jt})$.

- Number of parameters increases quadratically with J_g but not with J.

Multi-stage Budgeting (under AIDS) [2]

- The demand system at the **group stage** is:

$$\frac{y_{gt}}{y_t^*} = \alpha_g^{(2)} + \gamma_g^{(2)} \ \ln\left(\frac{y_t^*}{P_t^*}\right) + \sum_{g'=1}^{G} \beta_{g,g'}^{(2)} \ \ln(P_{gt})$$

$y_t^* =$ Total expenditure in the category (e.g., cereals);
$P_t^* =$ Price index for the category (e.g., cereals).

- Finally, at the top-stage, the **demand for the category** is:

$$\frac{y_t^*}{y_t} = \alpha^{(3)} + \gamma^{(3)} \ \ln(y_t) + \beta^{(3)} \ \ln(P_t^*)$$

Limitations

- Demand models in 'product space' have several practical limitations.

  1. **Representative consumer assumption.**

  2. **Too many parameters.**

  3. **Finding instruments for prices.**

  4. **Problems to predict demand of new varieties.**

[1] Representative consumer assumption

- Very unrealistic. Propensity to substitute between different products is very heterogeneous across consumers.
- Ignoring this heterogeneity can generate substantial biases.
- In principle, the model can be applied to consumer/household level data. However:
  - Household-level data is often not available for some products / industries.
  - At the lower-stage, observed household choices seem discrete (only one variety) and this is at odds with this "continuous choice" model.

[2] Too many parameters

- The number of parameters is $2J + \frac{J(J+1)}{2}$, i.e., J intercept parameters (α); J income elasticities (γ); and $\frac{J(J+1)}{2}$ free price elasticities (β).
- It is not possible to estimate demand systems for differentiated products with many varieties.
- For instance, demand system for car models. With $J = 100$, the $\#$parameters $= 5,250$.
- We need many thousands of observations (markets or/and time periods) to estimate this model. This type of data is typically not available.

[3] Finding instruments for prices

- Most applications of this class of models have ignored the potential endogeneity of prices.
- However, it is well known and simultaneity and endogeneity are potentially important issues in any demand estimation.
- The typical solution to this problem is using instrumental variables.
- In this model, the researcher needs at least as many instruments as prices, that is J.
- The ideal case is when we have information on production costs for each individual good. However, that information is very rarely available.

[4] Problems to predict demand of new varieties

- A problem that has received substantial attention is the prediction of the demand of a new product.
- Trajtenberg (1989), Hausman (1996), and Petrin (2002) are some prominent applications.
- In a demand system in product space, estimating the demand of a new good, say $J+1$, requires estimates of the parameters associated with that good: α_{J+1}, γ_{J+1} and $\{\beta_{J+1,j} : j = 1, 2, ..., J+1\}$.
- This makes it impossible to make counterfactual predictions, i.e., predict the demand of a product that has not been introduced in any market yet.
- It also limits the applicability of this model in cases where the new product has been introduced very recently or in very few markets, because we may not have enough data to estimate these parameters.

An Application: Hausman (1996) on cereals

- Hausman (1996) presents an application of demand in product space to an industry with many varieties: ready-to eat (RTE) cereals in US.
- This industry has been characterized by the proliferation of many varieties. Period 1980-92: 190 new brands were added to the pool of existing 160 brands.
- He deals with the limitations mentioned above by using:
(a) Multi-stage budgeting (and focusing on most popular varieties);
(b) Data from many periods (weekly data) and multiple geographic markets (cities), and assuming that parameters are constant across weeks-markets (up to fixed effects in the intercepts).
(c) Exploiting assumptions on the geographic structure of demand/supply shocks to generate instruments for prices.
(d) Evaluates the introduction of a new brand (Cheerios).

Hausman (1996) on cereals: Data

- Supermarket scanner data: period 1990-1992.
- 137 weeks ($T = 137$); 7 geographic markets ($M = 7$) or standard metropolitan statistical areas (SMSAs), including Boston, Chicago, Detroit, Los Angeles, New York City, Philadelphia, and San Francisco.
- Though the data includes information from hundred of brands, the model and the estimation concentrates in 20 brands classified in three segments: adult (7 brands), child (4 brands), and family (9 brands).
- $\{p_{jmt}, q_{jmt} : j = 1, 2, ..., 20;\ m = 1, 2, ..., 7;\ t = 1, 2, ..., 137\}$.
- Quantities are measured is physical units.
- There are not observable cost shifters.

Hausman (1996) on cereals: Model

- Almost-Ideal-Demand-System

$$w_{jmt} = \alpha^1_{jm} + \alpha^2_t + \gamma_j \ \ln\left(\frac{y_{gmt}}{P_{gmt}}\right) + \sum_{k \in J_g} \beta_{jk} \ \ln(p_{kmt}) + \varepsilon_{jmt}$$

- The terms α^1_{jm} and α^2_t represent brand-city and time "fixed effects".
- Suppose that the supply (pricing equation) is:

$$\ln(p_{jmt}) = \delta_j \ c_{jt} + \tau_{jm} + u_{jmt}$$

- c_{jt} represents a common cost shifter (unobservable to the researcher);

- τ_{jm} is city-brand fixed effect that captures differences in transportation costs;

- u_{jmt} captures potential response of prices to local demand shocks.

Hausman (1996) on cereals: Instruments

- The identification assumption is that demand shocks are not (spatially) correlated across markets: for any pair of markets $m \neq m$ it is assumed that:

$$E\left(u_{jmt}\ u_{km't}\right) = 0 \qquad \text{for any } j, k$$

- After controlling for brand-city fixed effects, all the correlation between prices at different locations comes from correlation in costs, and not from spatial correlation in demand shocks.

- Under these assumptions we can use average prices in other local markets, $\overline{P}_{j(-m)t}$, as instruments, where:

$$\overline{P}_{j(-m)t} = \frac{1}{M-1} \sum_{m' \neq m} p_{jm't}$$

Hausman (1996): Value of cheerios

- Hausman uses the estimated demand system to evaluate the value of a new varity that was introduced during this period: apple-cinamon cheerios (ACC).
- He first obtains the value of the price ACC that makes the demand of this product equal to zero. He obtains a virtiual price of \$7.14 (double the actual observed price \$3.5).
- Given this price, he calculates the consumer surplus (alternatively the CV or the EV).
- He obtains estimated welfare gains of \$32,268 per city and weekly average with an standard error of \$3,384.
- Aggregated at the level of US and annually, the consumer-welfare gain is \$78.1 million (or \$0.31 per person per year) which is a sizable amount of consumer's surplus.

Demand in characteristics space: Basic assumptions

- The basic assumptions are:
- (1) A product, say a laptop computer, can be described in terms of a bundle of physical characteristics: e.g., CPU speed, memory, screen size, etc. These characteristics determine a *variety* of the product.
- (2) Consumers have preferences on bundles of characteristics of products, not on the products per se.
- (3) A product has J different varieties and each consumer buys at most one variety of the product per period, i.e., the varieties are substitutes.

Model: Products

- We index varieties by $j \in \{1, 2, ..., J\}$.
- We can distinguish two sets of product characteristics.
- Characteristics observable and measurable to the researcher:

$$\mathbf{x}_j \equiv (X_{1j}, X_{2j}, ..., X_{Kj})$$

  X_{kj} represents the "amount" of attribute k in brand j.

- Example: Laptops: X_{1j} = CPU speed; X_{2j} = RAM memory; X_{3j} = hard disk memory; X_{4j} = weight; X_{5j} = screen size; X_{6j} = dummy 'Intel inside'; etc.
- Other characteristics are not observable to the researcher but they known and valuable to consumers:

$$\widetilde{\boldsymbol{\xi}}_j = (\xi_{1j}, \xi_{2j}, ...)$$

Model: Consumers

- We index households by $h \in \{1, 2, ..., H\}$ where H represents the number of households in the market.

- A household has preferences defined over bundles of attributes:

$$V_h(\mathbf{x}, \widetilde{\boldsymbol{\xi}})$$

$V_h(\mathbf{x}, \widetilde{\boldsymbol{\xi}})$ is the utility for consumer h of bundle of attributes $(\mathbf{x}, \widetilde{\boldsymbol{\xi}})$.

- Note that the utility function $V_h(\mathbf{x}, \widetilde{\boldsymbol{\xi}})$ is defined over any possible bundle of attributes whether or not this bundle exists in the market.

- For a product j in the market, this utility is $V_{hj} = V_h(\mathbf{x}_j, \widetilde{\boldsymbol{\xi}}_j)$.

Model: Consumer heterogeneity

- Consumers have different preferences and income, y_h.
- Consumer heterogeneity in preferences can be represented in terms of a vector of consumer attributes $\boldsymbol{v}_h$ that may be unobservable to the researcher.
- We can write the utility of consumer h as:

$$U_h = u(c, \boldsymbol{v}_h) + V(\mathbf{x}, \widetilde{\boldsymbol{\xi}}, \boldsymbol{v}_h)$$

- There is continuum of consumers with measure H, and $\boldsymbol{v}_h$ has a density function f_v in the market.

Model: Consumer decision

- Each consumer buys at most one variety of the product (per period).
- Given his income, y_h, and the vector of product prices $\mathbf{p} = (p_1, p_2, ..., p_J)$, a consumer decides which variety to buy, if any.
- Let $d_{hj} \in \{0, 1\}$ be the indicator of the event "consumer h buys product j". A consumer decision problem is, $d_{hj} = 1$ iff:

$$\max_{\{\mathbf{d}_h\}} \quad j = \arg\max_k \; u(y_h - p_k, \boldsymbol{\upsilon}_h) + V(\mathbf{x}_k, \widetilde{\boldsymbol{\xi}}_k; \boldsymbol{\upsilon}_h)$$

- Solution: Consumer unit demands: $d_j^*(\mathbf{x}, \mathbf{p}, y_h, \boldsymbol{\upsilon}_h)$.

Model: Aggregate demand

- Given consumers demands $d_j^*(\mathbf{p}, y_h, \boldsymbol{v}_h)$ and the joint density function $f(\boldsymbol{v}_h, y_h)$, we can obtain the aggregate demand functions:

$$q_j(\mathbf{x}, \mathbf{p}, f) = \int d_j^*(\mathbf{x}, \mathbf{p}, y_h, \boldsymbol{v}_h) \ , \beta) \ f(\boldsymbol{v}_h, y_h) \ d\boldsymbol{v}_h \ d\mathbf{y}_h$$

- And the market shares:

$$s_j(\mathbf{x}, \mathbf{p}, f) \equiv \frac{q_j(\mathbf{x}, \mathbf{p}, f)}{H}$$

Example: Logit model of product differentiation

- Suppose that:

$$V(\mathbf{x}_j, \widetilde{\boldsymbol{\xi}}_j; \boldsymbol{v}_h) = \mathbf{x}_j \ \beta + \xi_j + \varepsilon_{hj}$$

where ε's are $i.i.d.$ Extreme Value Type 1.

- And

$$u(C; \boldsymbol{v}_h) = \alpha \ C$$

- Then,

$$U_{hj} = -\alpha \ p_j + \mathbf{x}_j \ \beta + \xi_j + \varepsilon_{hj}$$

- And

$$q_j = H \ \frac{\exp\left\{-\alpha \ p_j + \mathbf{x}_j \ \beta + \xi_j\right\}}{1 + \sum_{k=1}^{J} \exp\left\{-\alpha \ p_k + \mathbf{x}_k \ \beta + \xi_k\right\}}$$

Example: Random coefficients

- Suppose that the utilities $V(\mathbf{x}_j, \widetilde{\boldsymbol{\xi}}_j; \boldsymbol{v}_h)$ and $u(C; \boldsymbol{v}_h)$ are linear in parameters, but these parameters are household specific. That is:

$$U_{hj} = -\alpha_h \ p_j + \mathbf{x}_j \ \boldsymbol{\beta}_h + \xi_j + \varepsilon_{hj}$$

- ε's are $i.i.d.$ Extreme Value Type 1, but

$$\begin{bmatrix} \alpha_h \\ \boldsymbol{\beta}_h \end{bmatrix} = \begin{bmatrix} \alpha \\ \boldsymbol{\beta} \end{bmatrix} + v_h$$

and

$$v_h \sim i.i.d. \ N(0, \Sigma)$$

Example: Random coefficients (2)

- Then, we can write utilities as:

$$U_{hj} = -\alpha \ p_j + \mathbf{x}_j \ \boldsymbol{\beta} + \xi_j + \widetilde{v}_{hj} + \varepsilon_{hj}$$

- where

$$\widetilde{v}_{hj} = -v_h^{\alpha} \ p_j + v_h^{\beta_1} \ X_{1j} + ... + v_h^{\beta_K} \ X_{Kj}$$

that has an heteroscedastic normal distribution.

- Then,

$$q_j = H \int \frac{\exp\left\{-\alpha \ p_j + \mathbf{x}_j \ \beta + \xi_j + \widetilde{v}_{hj}\right\}}{1 + \sum_{k=1}^{J} \exp\left\{-\alpha \ p_k + \mathbf{x}_k \ \beta + \xi_k + \widetilde{v}_{hk}\right\}} \ f\left(\widetilde{\mathbf{v}}_h | \Sigma\right) d\widetilde{\mathbf{v}}_h$$

Unobserved consumer heterogeneity

- In general, the more flexible is the structure of the unobserved consumer heterogeneity, the more flexible and realistic can be the elasticities of substitution between products that the model can generate.
- The logit model imposes very strong, and typically unrealistic, restrictions on demand elasticities.
- The random coefficients model generate much more flexible elasticities.

Estimation: Data & Parameters

- Suppose that the researcher has a dataset from a **single market** at only **one period** but for a product with many varieties: $M = T = 1$ but large J (e.g., 100 varieties).

- The researcher observes:

$$Data = \{q_j,\ X_j,\ p_j : j = 1, 2, ..., J\}$$

- Given these data, the researcher is interested in the estimation of the parameters of the demand system:

$$\theta = \{\alpha,\ \beta,\ \Sigma\}$$

- For the moment, we assume that market size H is known to the researcher. But it can be also estimated as a parameter.

- For the asymptotic properties of the estimators, we consider that $J \longrightarrow \infty$.

Estimation: Econometric Model

- The model is:

$$s_j = \frac{q_j}{H} = \sigma_j\left(\mathbf{x}, \mathbf{p}, \boldsymbol{\xi}; \theta\right)$$

- Unobserved characteristics $\boldsymbol{\xi}$ are correlated with $\mathbf{p}$ (endogeneity).
- Dealing with endogeneity in nonlinear models is complicated. Without further restrictions, we need full MLE: an specification of the model of $\mathbf{p}$ and a parametric specification of the distribution of $\boldsymbol{\xi}$.
- BLP show that there is a general class of models (BLP models) with an **invertibility property**.
- This property implies that we can represent the model using a equation where the unobservables $\boldsymbol{\xi}$ enter additively and linearly, and then we can estimate these equations using GMM.

Estimation: BLP invertibility

- Suppose that the utility function is:

$$U_{hj} = -\alpha \; p_j + \mathbf{x}_j \; \boldsymbol{\beta} + \xi_j + \widetilde{v}_{hj} + \varepsilon_{hj}$$

- Define:

$$\delta_j \equiv -\alpha \; p_j + \mathbf{x}_j \; \boldsymbol{\beta} + \xi_j$$

  such that $U_{hj} = \delta_j + \widetilde{v}_{hj} + \varepsilon_{hj}$. The term δ_j can be interpreted as the objective value of product j.

- Let δ be the vector of $\delta_j s$. Then,

$$s_j = \sigma_j \left(\delta; \, \mathbf{x}, \mathbf{p}, \Sigma \right)$$

- In vector form,

$$\mathbf{s} = \boldsymbol{\sigma} \left(\delta; \, \mathbf{x}, \mathbf{p}, \Sigma \right)$$

Estimation: BLP invertibility

- **Berry (1994) Proposition.** Under some regularity conditions (more later) the system

$$\mathbf{s} = \sigma\left(\delta; \mathbf{x}, \mathbf{p}, \Sigma\right)$$

is invertible in δ such that there is an inverse function $\sigma^{-1}(.)$ and:

$$\delta = \sigma^{-1}\left(\mathbf{s}; \mathbf{x}, \mathbf{p}, \Sigma\right)$$

- For a particular product:

$$\delta_j = \sigma_j^{-1}\left(\mathbf{s}; \mathbf{x}, \mathbf{p}, \Sigma\right)$$

- The form of the inverse mapping σ^{-1} depends on the PDF $f_{\widetilde{v}}$.

Example: Logit model

- In the logit model:

$$s_j = \frac{\exp\{\delta_j\}}{1 + \sum_{k=1}^{J} \exp\{\delta_k\}}$$

- Let s_0 be the market share of the "outside good" such that, $s_0 = 1 - \sum_{k=1}^{J} s_k$.

- Then,

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

- That is,

$$\sigma_j^{-1}(\mathbf{s}; \mathbf{x}, \mathbf{p}, \Sigma) = \ln(s_j) - \ln(s_0)$$

and we have a closed form expression for the inverse mapping σ_j^{-1}.

Example: Nested Logit model

- No RCs. G groups of product, indexed by g. The ε_{hj} "taste" variable has the structure of a nested logit: $\varepsilon_{hj} = \lambda\ \varepsilon_{hg}^{(1)} + \varepsilon_{hj}^{(2)}$, where $\varepsilon_{hg}^{(1)}$ and $\varepsilon_{hj}^{(2)}$ are i.i.d. EV type 1 variables, and λ is a parameter.
- In the logit model:

$$s_j = \frac{\exp\left\{\lambda\ I_g\right\}}{\sum_{g'=0}^{G} \exp\left\{\lambda\ I_{g'}\right\}} \frac{\exp\left\{\delta_j\right\}}{\sum_{j\in J_g} \exp\left\{\delta_j\right\}}$$

- Then,

$$\ln\left(s_j\right) - \ln\left(s_0\right) - \lambda\ \ln\left(s_{j|g}\right) = \delta_j$$

- That is,

$$\sigma_j^{-1}\left(\mathbf{s};\ \mathbf{x}, \mathbf{p}, \Sigma\right) = \ln\left(s_j\right) - \ln\left(s_0\right) - \lambda\ \ln\left(s_{j|g}\right)$$

where $s_{j|g}$ is the within-group market share.

Numerical inversion: Fixed Point algorithm

- For the RC-Logit there is not a closed form expression for σ_j^{-1}.
- BLP propose a fixed point algorithm to compute the δ's.
- Consider the mapping from the demand system [omitting the other arguments $(\mathbf{x}, \mathbf{p}, \Sigma)$: $\mathbf{s} = \sigma(\delta)$, or at the product level $s_j = \sigma_j(\delta)$.
- BLP propose instead solving the fixed point mapping in δ:

$$\delta = \mathbf{F}(\delta, \mathbf{s})$$

where $\mathbf{F}(\delta, \mathbf{s}) = \{F_j(\delta, s_j) : j = 1, 2, ..., J\}$ and:

$$F_j(\delta, s_j) \equiv \delta_j + \ln(s_j) - \ln(\sigma_j(\delta))$$

Numerical inversion: Fixed Point algorithm [2]

- It is clear that $\delta_j = F_j(\delta, s_j)$ [fixed point] implies $\delta_j = \delta_j + \ln(s_j) - \ln(\sigma_j(\delta))$ and this implies $s_j = \sigma_j(\delta)$.

- BLP show that the mapping

$$F_j(\delta, s_j) \equiv \delta_j + \ln(s_j) - \ln(\sigma_j(\delta))$$

  is a contraction as long as the values of δ are not too small.

- Under this condition, the mapping has a unique fixed point and we can find it by using fixed point iteration algorithm.

- Iterative procedure: Start with initial δ^0. Iteration $R+1$, for every j:

$$\delta_j^{R+1} = \delta_j^R + \ln(s_j) - \ln\left(\sigma_j\left(\delta^R\right)\right)$$

- Iterate until convergence.

Numerical inversion: Fixed Point algorithm [3]

- Note that this "inversion" should be done for a given value of $(\mathbf{x}, \mathbf{p}, \Sigma)$.

- If we have multiple markets in the data, $\{\mathbf{x}_t, \mathbf{p}_t : t = 1, 2, ..., T\}$, we need to do this inversion for every market.

- Importantly, for every "trial value" of Σ in the (non-linear) estimation algorithm of Σ, we need to re-evaluate these fixed points.

GMM Estimation

- The model can be described as [with potentially multiple markets t]:

$$\xi_{jt} = \sigma_j^{-1}\left(\mathbf{s}_t, \mathbf{x}_t, \mathbf{p}_t, \Sigma\right) - \left[\alpha \ p_{jt} - \mathbf{x}_{jt} \ \boldsymbol{\beta}\right]$$

- We want to estimate $\theta = \{\alpha, \ \beta, \ \Sigma\}$ in this model.

- Suppose that we have a vector of instruments $\mathbf{Z}_{jt}$ such that:
  (1) $E(\xi_{jt} \mid \mathbf{Z}_{jt}) = 0$
  (2) $\dim(\mathbf{Z}_{jt}) \geq \dim(\theta)$
  (3) $E\left[\left(\frac{\partial\sigma_j^{-1}(\mathbf{s}_t, \mathbf{x}_t, \mathbf{p}_t, \Sigma)}{\partial\Sigma}, \ p_{jt}, \ \mathbf{x}_{jt}\right)' \left(\frac{\partial\sigma_j^{-1}(\mathbf{s}_t, \mathbf{x}_t, \mathbf{p}_t, \Sigma)}{\partial\Sigma}, \ p_{jt}, \ \mathbf{x}_{jt}\right) \mid \mathbf{Z}_{jt}\right]$ is non-singular.

- More on condition (3) later ...

GMM Estimation [2]

- Under conditions (1) to (3) we can estimate θ using GMM. Define:

$$m(\theta) = \frac{1}{JT} \sum_{j=1}^{J} \sum_{t=1}^{T} \mathbf{Z}_{jt} \left[\sigma_j^{-1} \left(\mathbf{s}_t, \mathbf{x}_t, \mathbf{p}_t, \Sigma \right) - \alpha \ p_{jt} + \mathbf{x}_{jt} \ \boldsymbol{\beta} \right]$$

- GMM estimator:

$$\widehat{\theta} = \arg\min_{\theta} \left[m(\theta)' \ \mathbf{W} \ m(\theta) \right]$$

- $m(\theta)$ is nonlinear in Σ. There is not a closed-form expression for $\widehat{\theta}$: we need a optimization algorithm, e.g., Newton's method.
- Nested Fixed Point algorithm. Alternative methods: (a) MPEC (Dube, Fox, & Su, ECTA 2012); (b) Nested Pseudo Likelihood (Lee & Seo, RAND 2015).

BLP instruments

- Exogenous characteristics of other products.
- For instance,

$$\mathbf{z}_{jt} = \frac{1}{J-1} \sum_{k \neq j} \mathbf{x}_{kt}$$

or the average $\mathbf{x}_{kt}$ for the "n-closest-neighbors" to $\mathbf{x}_{jt}$.

Hausman-Nevo instruments

- Prices in other geographic markets, after controlling for product fixed effects.
- Suppose t represents geographic markets and the T markets belong to R regions.
- Suppose that: $\xi_{jt} = \xi_j^{(1)} + \xi_t^{(2)} + \xi_{jt}^{(3)}$, and $\xi_{jt}^{(3)}$ is NOT spatially correlated.
- Supply shocks are spatially correlated between local markets within the same region, but demand shocks are not.
- Define: $Z_{jt} = \frac{1}{T_R - 1} \sum_{t' \in R_{-t}} p_{jt'}$.
- Under these assumptions, $E(\xi_{jt}^{(3)} \mid Z_{jt}) = 0$, and Z_{jt} is correlated with p_{jt}.

Arellano-Bond instruments

- Suppose t represents time and $\xi_{jt} = \xi_j^{(1)} + \xi_t^{(2)} + \xi_{jt}^{(3)}$, and $\xi_{jt}^{(3)}$ is NOT serially correlated.
- Suppose that supply shocks are serially correlated.
- Define $\mathbf{Z}_{jt} = \{s_{jt-2},\ p_{jt-2}\}$.
- Under these assumptions, $E(\Delta\xi_{jt}^{(3)} \mid \mathbf{Z}_{jt}) = 0$, and Z_{jt} is correlated with p_{jt}.

3. Valuation of new products

Valuation of new products

- Consider an individual with preference parameters $(\alpha_h, \beta_h, \varepsilon_h)$ facing a set of products $\mathcal{J}$ with vector of prices $\mathbf{p}$.
- The indirect utility function is defined as (income effects are assumed away because linearity):

$$v(\mathbf{p}, \alpha_h, \beta_h, \varepsilon_h) = \max_{j \in \mathcal{J}} \left[-\alpha_h p_j + \mathbf{x}_j \beta_h + \xi_j + \varepsilon_{hj} \right]$$

- Following McFadden (1981) and Small & Rosen (1981), we define the **money-metric welfare function**:

$$\frac{1}{\alpha_h} \; v(\mathbf{p}, \alpha_h, \beta_h, \varepsilon_h)$$

- And its aggregate version:

$$W(\mathbf{p}) = \int \frac{1}{\alpha_h} \; v(\mathbf{p}, \alpha_h, \beta_h, \varepsilon_h) \; dF(\alpha_h, \beta_h, \varepsilon_h)$$

Valuation of new products [2]

- For the random coefficients logit model:

$$W(\mathbf{p}) = \int \frac{1}{\alpha_h} \ln \left[\sum_{j=0}^{J} \exp \left\{ -\alpha_h p_j + \mathbf{x}_j \beta_h + \xi_j \right\} \right] \, dF(\alpha_h, \beta_h)$$

- We can include $\mathbf{x}$ and $\mathcal{J}$ as explicit arguments of the welfare function: $W(\mathbf{p}, \mathbf{x}, \mathcal{J})$.

- We can use $W(.)$ to measure the welfare effects of a change in:
  - Prices: $W(\mathbf{p}^1, \mathbf{x}, \mathcal{J}) - W(\mathbf{p}^0, \mathbf{x}, \mathcal{J})$
  - Products characteristics: $W(\mathbf{p}, \mathbf{x}^1, \mathcal{J}) - W(\mathbf{p}, \mathbf{x}^0, \mathcal{J})$
  - Set of products: $W(\mathbf{p}, \mathbf{x}, \mathcal{J}^1) - W(\mathbf{p}, \mathbf{x}, \mathcal{J}^0)$

Some limitations

- [1] Problems to evaluate **radical innovations** with new types of characteristics.
- [2] **Logit errors**. There is very limited "crowding" of products, i.e., $\lim_{J \to \infty} W(\mathcal{J}) = \infty$ (see Ackerberg and Rysman, RAND 2005).
- [3] **Outside alternative**. Unobserved "qualities" ξ_{jt} are relative to the outside alternative. Suppose there is quality improvements in the outside alternative. We will underestimate the welfare improvements in this industry.

Trajtenberg (1989) on computed tomography scanners

- Computed tomography (CT) scanners is considered a key innovation in imaging diagnosis in medicine during the 1970s.
- The first was installed in US in 1973, and soon after 20 firms entered in this market with different varieties, General Electric being the leader.
- Clients are hospitals.
- Three characteristics are key for the quality: scan time; image quality; and reconstruction time.

Trajtenberg: Data

- Period: 1973-1981.
  - 55 products.
  - Product characteristics (price, scan speed, resolution, reconstruction speed) and sales in US.
  - identity and attributes of the buying hospital.
- Hospital-year level data: the dependent variable is the product choice of hospital h at year t.

Trajtenberg: Estimated model

- Nested logit at the hospital-year level: nests are "head" and "body" scanner.
- Utility: Quadratic in the three product attributed (other than price).
- Estimated elasticity of substitution between the two "nests" is very close to zero, i.e., two different products.
- Estimated $\alpha' s$ are significant but with the wrong sign for "body" [Not accounting for endogeneity of prices].

Trajtenberg: Welfare effects

- The counterfactual is eliminating all CT scanner products (keeping only the outside product
- The estimated welfare effect of CT scanners during this period is **$16 million** (of 1982).
- Using data of firms' R&D investment, he obtains a social rate of return of **270%**.

Petrin (JPE, 2002) on minivans

- Evaluate consumer welfare gains from the introduction of a new type of car, the minivan.
- Estimation of a BLP demand system of automobiles. Combine market level and micro moments.
- Observing average family size conditional on the purchase of a minivan and asking the model to match this average helps to identify parameters that capture consumer taste for the characteristics of minivans.

Petrin on minivans [2]

- In 1984, Chrysler introduced the Dodge caravan (its minivan). It was an immediate success.
- GM and Ford responded quickly introducing in 1985 their own.
- By 1998, there were 6 firms selling a total of 13 different minivans, Chrysler being the leader (44%).

Petrin on minivans: Data

- Period: 1981-1993. $J = 2407$. Product-year panel.
- Variables: Quantity sold; price, acceleration, dimensions, drive type, fuel efficiency, a measure of luxury.
- Consumer expenditure survey (CEX).links demographics of purchasers of new vehicles to the vehicles they purchase.
- In the CEX, we observe 2,660 new vehicle purchases over the period and sample. Used to estimate the probabilities of new vehicle purchases for different income groups.
- Observed purchases of minivans (120), station wagons (63), SUVs (131), and full-size vans (23). Used to estimate average family size and age of purchasers of each of these vehicle types.

Petrin on minivans: Market shares

Family Vehicle Sales as a Percentage of Total Vehicle Sales: U.S. Automobile Market, 1981–93

Year	Minivans (1)	Station Wagons (2)	Sport-Utilities (3)	Full-Size Vans (4)	Minivans and Station Wagons (5)	U.S. Auto Sales (Millions) (6)
1981	.00	10.51	.58	.82	10.51	7.58
1982	.00	10.27	.79	1.17	10.27	7.05
1983	.00	10.32	3.51	1.04	10.32	8.48
1984	1.58	8.90	5.51	1.20	10.48	10.66
1985	2.32	7.33	6.11	1.05	9.65	11.87
1986	3.63	6.70	5.73	.85	10.43	12.21
1987	4.86	6.47	6.44	.73	11.33	11.21
1988	5.97	5.14	7.18	.69	11.11	11.76
1989	6.45	4.13	7.47	.61	10.58	11.06
1990	7.95	3.59	7.78	.27	11.54	10.51
1991	8.29	3.05	7.80	.29	11.34	9.75
1992	8.77	3.07	9.33	.39	11.84	10.12
1993	9.93	3.02	11.66	.29	12.95	10.71

Petrin on minivans: Estimates

TABLE 4
PARAMETER ESTIMATES FOR THE DEMAND-SIDE EQUATION

Variable	OLS Logit (1)	Instrumental Variable Logit (2)	Random Coefficients (3)	Random Coefficients and Microdata (4)
	A. Price Coefficients (α's)			
α_1	.07	.13	4.92	7.52
	(.01)**	(.01)**	(9.78)	(1.24)**
α_2			11.89	31.13
			(21.41)	(4.07)**
α_3			37.92	34.49
			(18.64)**	(2.56)**

Petrin on minivans: Estimates

	B. Base Coefficients (β's)			
Constant	−10.03	−10.04	−12.74	−15.67
	(.32)**	(.34)**	(5.65)**	(4.39)**
Horsepower/weight	1.48	3.78	3.40	−2.83
	(.34)**	(.44)**	(39.79)	(8.16)
Size	3.17	3.25	4.60	4.80
	(.26)**	(.27)**	(24.64)	(3.57)*
Air conditioning standard	−.20	.21	−1.97	3.88
	(.06)**	(.08)**	(2.23)	(2.21)*
Miles/dollar	.18	.05	−.54	−15.79
	(.06)**	(.07)	(3.40)	(.87)**
Front wheel drive	.32	.15	−5.24	−12.32
	(.05)**	(.06)**	(3.09)	(2.36)**
Minivan	.09	−.10	−4.34	−5.65
	(.14)	(.15)	(13.16)	(.68)**
Station wagon	−1.12	−1.12	−20.52	−1.31
	(.06)**	(.07)**	(36.17)	(.36)**
Sport-utility	−.41	−.61	−3.10	−4.38
	(.09)**	(.10)**	(10.76)	(.41)**
Full-size van	−1.73	−1.89	−28.54	−5.26
	(.16)**	(.17)**	(235.51)	(1.30)**
% change GNP	.03	.03	.08	.24
	(.01)**	(.01)**	(.02)**	(.02)**

Petrin on minivans: Random coef.

RANDOM COEFFICIENT PARAMETER ESTIMATES

	RANDOM COEFFICIENTS (γ's)	
VARIABLE	Uses No Microdata (1)	Uses CEX Microdata (2)
Constant	1.46	3.23
	(.87)*	(.72)**
Horsepower/weight	.10	4.43
	(14.15)	(1.60)**
Size	.14	.46
	(8.60)	(1.07)
Air conditioning standard	.95	.01
	(.55)*	(.78)
Miles/dollar	.04	2.58
	(1.22)	(.14)**
Front wheel drive	1.61	4.42
	(.78)**	(.79)**
γ_{mi}	.97	.57
	(2.62)	(.10)**
γ_{sw}	3.43	.28
	(5.39)	(.09)**
γ_{su}	.59	.31
	(2.84)	(.09)**
γ_{pv}	4.24	.42
	(32.23)	(.21)**

Petrin on minivans: Marginal costs

TABLE 6
PARAMETER ESTIMATES FOR THE COST SIDE
Dependent Variable: Estimated (Log of) Marginal Cost

Variable (τ's)	Parameter Estimate	Standard Error
Constant	1.50	.08
ln(horse power/weight)	.84	.03
ln(weight)	1.28	.04
ln(MPG)	.23	.04
Air conditioning standard	.24	.01
Front wheel drive	.01	.01
Trend	−.01	.01
Japan	.12	.01
Japan × trend	−.01	.01
Europe	.47	.03
Europe × trend	−.01	.01
ln(q)	−.05	.01

Petrin on minivans: Welfare evaluation

- Counterfactual: no minivans.
- It takes into account the counterfactual equilibrium prices without minivans (based on estimates of demand and marginal costs).
- The introduction of minivans (particularly, Dodge caravan) had an important negative effect on prices of many substitutes that were top-selling vehicles in the large-sedan and wagon segments of the market.
- There were also some price increases due to cannibalization of own products.

Petrin: Price changes with / without minivans

Equilibrium Prices with and without the Minivan, 1984:
1982–84 CPI-Adjusted Dollars

	Price			
	With Minivan	Without Minivan	ΔPrice	% ΔPrice
	A. Largest Price Decreases on Entry			
GM Oldsmobile Toronado (large sedan)	15,502	15,643	−141	.90
GM Buick Riviera (large sedan)	15,379	15,519	−139	.89
GM Buick Electra (large sedan)	12,843	12,978	−135	1.04
GM Chevrolet Celebrity (station wagon)	8,304	8,431	−127	1.51
Ford Cadillac Eldorado (large sedan)	19,578	19,704	−126	.64
Ford Cadillac Seville (large sedan)	21,625	21,749	−125	.57
GM Pontiac 6000 (station wagon)	9,273	9,397	−123	1.31
GM Oldsmobile Ciera (station wagon)	9,591	9,714	−123	1.27
GM Buick Century (station wagon)	8,935	9,056	−121	1.34
GM Oldsmobile Firenza (station wagon)	7,595	7,699	−104	1.35
	B. Largest Price Increases on Entry			
Chrysler LeBaron (station wagon)	9,869	9,572	297	3.10
Volkswagen Quattro (station wagon)	13,263	13,079	184	1.41
Chrysler (Dodge) Aries K (station wagon)	7,829	7,659	170	2.22
AMC Eagle (station wagon)	10,178	10,069	109	1.08

Petrin: Welfare evaluation [2]

- The mean per capita Compensated Variation for minivans is $1247.
- But the estimates are substantially larger with the other methods (OLS logit; IV logit: IV BLP without micro moments). This is mainly because these methods under-estimate the marginal utility of income.
- Decomposition of the welfare gains in the contribution of:
  - x_j and ξ_j: $851
  - logit ε_{hj}: $396
- Other methods imply very implausible contribution of the idiosyncratic ε's.

Petrin: Welfare gains

Average Compensating Variation Conditional on Minivan Purchase, 1984: 1982–84 CPI-Adjusted Dollars

	OLS Logit	Instrumental Variable Logit	Random Coefficients	Random Coefficients and Microdata
Compensating variation:				
Median	9,573	5,130	1,217	783
Mean	13,652	7,414	3,171	1,247
Welfare change from difference in:				
Observed characteristics ($\delta_j + \mu_{ij}$)	−81,469	−44,249	−820	851
Logit Error (ϵ_{ij})	95,121	51,663	3,991	396
Income of minivan purchasers:				
Estimate from model	23,728	23,728	99,018	36,091
Difference from actual (CEX)	−15,748	−15,748	59,542	−3,385

Ackerberg & Rysman (RAND, 2005)

- The Logit errors can have unrealistic implications on the evaluation of welfare gains.
- Because these errors, welfare increases unboundedly (though concavely) with J. No crowding.
  e.g., identical products except for the ε's: $W = \ln(\sum_{j=0}^{J} \exp\{\delta\}) = \delta \ln(J+1)$.
- Though the BLP or RC-Logit limits the influence of the logit errors, it is still subject to this problem.
- Ackerberg & Rysman (2005) propose a simple but interesting modification of the logit model that can account for this problem.

Ackerberg & Rysman [2]

- Consider a variation of the BLP model where the dispersion of the logit errors depends on the number of products in the market. For $j > 0$:

$$U_{hj} = -\alpha_h p_j + \mathbf{x}_j \beta_h + \xi_j + \sigma(J)\ \varepsilon_{hj}$$

  $\sigma(J)$ is strictly decreasing in J and it goes to 0 as J goes to ∞.

- As J increases, the differentiation from the $\varepsilon' s$ becomes less and less important.

- Function $\sigma(J)$ can be parameterized and its parameters estimated together with the rest of the model.

- Though Ackerberg & Rysman consider this approach, they favor a similar approach that is simpler to implement.

Ackerberg & Rysman [3]

- Instead, the consider the model. For $j > 0$:

$$U_{hj} = -\alpha_h p_j + \mathbf{x}_j \beta_h + \xi_j + f(J, \gamma) + \varepsilon_{hj}$$

where $f(J, \gamma)$ is a decreasing function of J parameterized by γ.

- For instance, $f(J, \gamma) = \gamma \ln(J)$.
- It can be also extended to a nested logit version. For group g: $f_g(J, \gamma) = \gamma_g \ln(J_g)$.
- The reasons for the specification $f(J, \gamma)$ instead of $\sigma(J, \gamma)$ is simplicity in estimation.

5. New goods with complementarity: Gentzkow (AER, 2007)

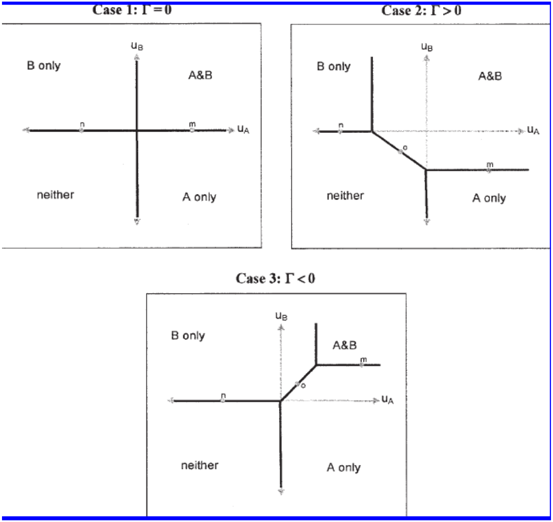
New goods with complementarity: Gentzkow (AER, 2007)

- The class of discrete choice demand models we have considered so far rule out complementarity between products.
- This is an important limitation in some relevant contexts:
  - Evaluating the merger between two firms producing complements, e.g., Pepsico and Frito-Lay
  - Evaluating the welfare effects of new products.
- Also, sometimes there are both substitution and complementarity effects: e.g., radio stations on recorded music; movie on the book novel; Uber and taxis.
- Gentzkow (2007) extends McFadenn / BLP framework to allow for complementarity, and studies the demand and welfare effect of online newspapers.

Gentzkow: Model

- Now consumers can choose bundles of products.
- Simple example: Two products A and B. Choice set: $\{0, A, B, AB\}$.
- Utilities: 0, u_A, u_B, and $u_{AB} = u_A + u_B + \Gamma$.
- Discrete choice: $P_j = \Pr(u_j = \max\{0, u_A, u_B, u_{AB}\})$.
- Quantities: $Q_A = P_A + P_{AB}$; $Q_B = P_B + P_{AB}$.
- Products A and B are: substitutes if $\partial Q_A / \partial p_B > 0$; complements if $\partial Q_A / \partial p_B < 0$.
- Complements / substitutes is closely related to the sign of Γ.

Gentzkow: Model



Gentzkow: Identification

- Suppose that: $u_{hA} = \beta_A - \alpha\ p_A + v_{hA}$; and $u_{hB} = \beta_B - \alpha\ p_B + v_{hB}$.
- Allowing for correlation between unobservables v_{hA} and v_{hB} is very important.
- Observing that frequent online readers are also frequent print readers might be evidence that the products in question are complementary, but it might also reflect that unobservable tastes for the goods are correlated.
- Suppose that (v_{hA}, v_{hB}) are standard normals with correlation ρ.
- The parameters of the model are: β_A, β_B, α, ρ, Γ.
- The researcher (with consumer level data) observes prices and bundles market shares: P_A, P_B, P_{AB}.

Gentzkow: Identification [2]

- Even with micro-level data on shares P_A, P_B, P_{AB}, the parameters $(\beta_A, \beta_B, \alpha, \rho, \Gamma)$ are NOT identified.
- Even if α is known, we have 3 data points and 4 parameters.
- Without further restrictions, a high value of P_{AB} can be explained by either high Γ or high ρ.
- But these two interpretations have very different economic and policy implications, including the evaluation of new products, e.g., including A when B was present.
- Gentzkow considers two sources of identification or additional moment conditions.

Gentzkow: Exclusion restrictions

- Suppose that there is an exogenous consumer characteristic (or vector) z that enters in consumer valuation of product A but not of product B.

$$\beta_A(z), \text{ but } \beta_B \text{ does not depend on } z$$

- For instance, if B is a print newspaper and A is its online version, z could be Internet access at work (at home could be more endogenous).

- Suppose z is binary for simplicity. Now, the data $[P_A(z), P_B(z), P_{AB}(z): z \in \{0, 1\}]$ can identify $\beta_A(0)$, $\beta_A(1)$, β_B, Γ, and ρ.

- Intuition: if $\Gamma > 0$ (complementaity), then $z = 1$ should increase $P_A(z)$ and $P_{AB}(z)$. Otherwise, if $\Gamma = 0$, then $z = 1$ should increase $P_A(z)$ but not $P_{AB}(z)$.

Gentzkow: Panel Data

- Suppose that we observe consumer choices at different periods of time. And suppose that:

$$v_{jht} = \eta_{jh} + \varepsilon_{jht}$$

- The time-invariant effects η_{Ah} and η_{Bh} are correlated with each other; but ε_{Aht} and ε_{Aht} are independent and iid over h, t.
- Preference parameters are assumed time invariant.
- Suppose that $T = 2$. We have 8 possible choice histories, 7 probabilities, and 4 parameters: β_A, β_B, Γ, and ρ.
- Identification intuition: if $\Gamma > 0$, changes over time in demand should be correlated between the two goods. If $\Gamma = 0$, changes over time should be uncorrelated between goods.

Gentzkow: Data

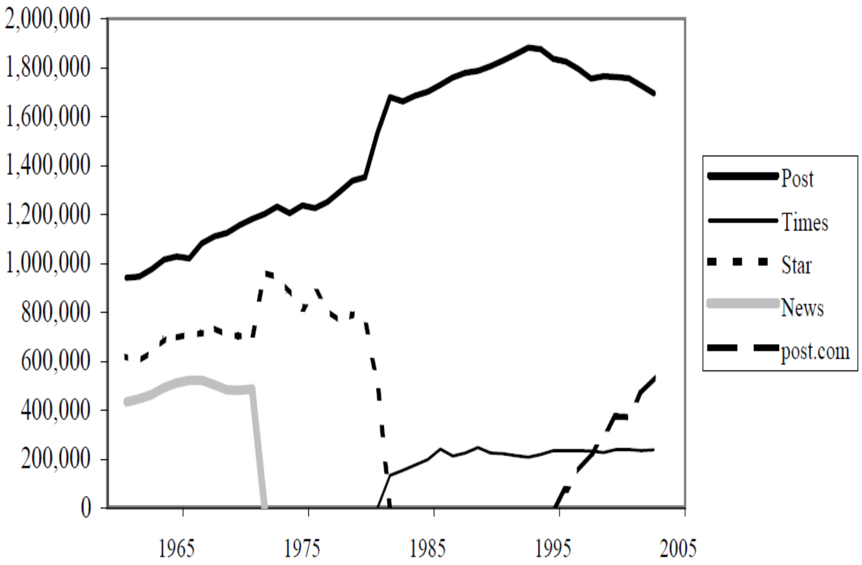
- Survey: 16,179 individuals in Washington DC, March-2000 and Feb. 2003.

- Information on individual and household characteristics, and readership of:
  - print local newspapers read over last week;
  - major local online newspapers over last week

- Two main local print newspapers: Times and Post. One main online newspaper: post.com.

- Three products: Times, Post, and post.com. Outside alternative being all the other local papers.

Gentzkow: Time-series readers



Gentzkow: Estimated Gammas

TABLE 6—PARAMETER ESTIMATES FROM FULL MODEL: OTHER

Interaction terms		*Excluded variables (coefficient in utility of post.com)*	
Post-post.com	−1.285**	Internet at work	1.357**
	(0.2307)		(0.180)
Post-*Times*	0.0809	Fast connection	0.146
	(0.2479)		(0.193)
post.com-*Times*	−1.231**	Use for education-related	0.361
	(0.4832)		(0.212)
Nonlinear parameters		Use for work	0.582**
τ	6.846**		(0.222)
	(0.5027)		
γ	0.0454**		
	(0.0179)		

Gentzkow: Effects of post.com

TABLE 8—IMPACT OF THE ONLINE EDITION ON DEMAND FOR PRINT

Case 1: Full model	
Cross-price derivative	8,358
	(1,436)
Change in print readership	−26,822
	(4,483)
Change in print profits	−$ 5,466,846
	(913,699)
Case 2: Model with observable characteristics only	
Cross-price derivative	−8,421
	(752)
Change in print readership	25,655
	(2,270)
Change in print profits	$ 5,229,009
	(462,771)
Case 3: Model with no heterogeneity	
Cross-price derivative	−16,143
	(702)
Change in print readership	51,897
	(2,254)
Change in print profits	$10,577,720
	(459,464)

Gentzkow: Main results

- Reduced-form OLS regressions and a structural model without heterogeneity suggest that the print and online editions of the Post are strong complements.
- According to those estimates, the addition of the post.com to the market increases profits from the Post print edition by $10.5 million per year.
- However, **properly accounting for consumer heterogeneity changes the conclusions substantially**.
- Estimates of the model with both observed and unobserved heterogeneity show that the print and online editions are **significant substitutes**.

Gentzkow: Main results [2]

- Raising the price of the Post by \$0.10 would increase post.com readership by about 2%.
- Removing the post.com from the market entirely would increase readership of the Post by 27,000 readers per day, or 1.5%.
- The estimated \$33.2 million of revenue generated by the post.com comes at a cost of about \$5.5 million in lost Post readership.
- For consumers, the online edition generated a per-reader surplus of \$0.30 per day, implying a total welfare gain of \$45 million per year.
- Reduced-form OLS regressions and a structural model without heterogeneity suggest that the print and online editions of the Post are strong complements.