Identification and Estimation of Demand for Differentiated Products

Jean-François Houde
Cornell University & NBER

September 24, 2016
Starting Point: The Characteristic Approach

- **Ultimate goal:** Measure the elasticity of substitution between goods.
- **Why?**
  - Measure the degree of market power
  - Predict the effect of proposed mergers
  - Evaluate the value of new goods
  - ...

- **Natural starting point:** Linear inverse demand

\[ p_{jt} = \alpha_0 + \sum_{k \in \mathcal{J}} \alpha_{j,k} q_{kt} + \epsilon_{jt} \]

- **Curse of dimensionality:** Even with exogenous variation in \( q \)'s, the number of parameters to estimate grows with the number of products.
- **Solution:** The characteristic approach (Lancaster 1966)
  - Each product is defined as a bundle of characteristics: \( x_j = \{x_1, \ldots, x_k\} \)
  - Consumers assigned values to each characteristics, and choose the product maximize their utility given their budget.
  - Demand for product \( j \) is function of the distribution of valuations for each attribute (finite dimension).
Representative Agent Models

- **Almost Ideal Demand System** (Deaton and Muellbauer 1980)
  - **Assumption:** Products are grouped in $K$ categories (i.e. $x$'s), and utility is separable in the sub-utility of the categories.
  - **Multi-stage budgeting:** Representative consumer allocates total expenditures in category and sub-categories. At each stage, only the prices (or price indices) of members of the group matter.
  - Typically leads to log-linear category demand systems
  - **Examples:**
    - Hausman, Leonard, and Zona (1994): The effect of mergers in the market for beers (i.e. categories = segments)

  - Semi-parametric representative-agent demand
  - The degree of substitutability between products is modeled using flexible functions of the distance between products.
  - Examples: Physical distance, alcohol content, calories, etc.
Discrete Choice Models

- **Two classes of models**
  - *Local:* Each product has only few close substitutes
  - *Global:* Each product is equally substitutable with all others

- **Global competition:** Logit (McFadden 1974, Anderson et al. 1992)
  - Indirect utility:
    \[
    u_{ij} = \delta_j - p_j + \epsilon_{ij}
    \]
  - Demand (shares):
    \[
    s_j = \frac{\exp((\delta_j - p_j)/\sigma_\epsilon)}{\sum_{j' \in J} \exp((\delta_{j'} - p_{j'})/\sigma_\epsilon)}
    \]
  - Elasticity of substitution:
    \[
    \eta_{jk} = \left( \frac{1}{\sigma_\epsilon} s_j s_k \right) \frac{p_k}{s_j} = \frac{1}{\sigma_\epsilon} p_k s_j
    \]
  - Two key properties: (i) Independence of Irrelevance Alternatives (IIA), and (ii) monopolistic competition in the limit (i.e. constant markup).
Local competition: Vertical or horizontal differentiation?

- **Definition:**
  - Horizontal differentiation: Consumers have different valuations of products available (even with equal prices)
  - Vertical differentiation: Consumers would choose the same product if prices were equal.

- **Horizontal Example:** Linear city model (Hotelling 1929)
  - Preferences for product $j = 1, 2$:
    \[ u_{ij} = \delta - p_j - \lambda |l_i - x_j| \]
    where $l_i \sim U(0,1)$ is the location (or taste) of consumer $i$, and $x_j \in (0,1)$ is the “address” of products (i.e. characteristic) and $x_2 > x_1$.
  - Demand:
    
    Type $\bar{l} \in [x_1, x_2]$:
    \[ \delta - p_1 - \lambda (\bar{l} - x_1) = \delta - p_2 - \lambda (x_2 - \bar{l}) \]
    \[ \bar{l} = \frac{(p_2 - p_1) + \lambda (x_2 - x_1)}{2} \]
    \[ s_1 = \int_0^{\bar{l}} f(l_i)dl_i = \frac{\lambda (x_2 - x_1) - p_1 + p_2}{2} \]
Local competition: Vertical or horizontal differentiation?

- **Vertical Example:** Quality ladder (Shaked and Sutton 1982, Bresnahan 1987)
  - Preferences for product $j = 1, 2$:
    \[
    u_{ij} = \theta_i x_j - p_j
    \]
    where $l_i \sim U(0, \lambda)$ is the location (or income) of consumer $i$, and $x_j \in (0, 1)$ is the “quality” of products (i.e. characteristic) and $x_2 > x_1$.
  - Demand:
    \[
    \text{Type } \bar{\theta}: \quad \bar{\theta} x_1 - p_1 = \bar{\theta} x_2 - p_2
    \]
    \[
    s_1 = \int_{0}^{\bar{\theta}} f(\theta_i) d\theta_i = \bar{\theta} / \lambda = \frac{1}{\lambda} \frac{p_2 - p_1}{x_2 - x_1}
    \]
Local competition: Vertical or horizontal differentiation?

- In both cases the elasticity of substitution is positive **only** for neighboring products
  - Spatial differentiation:
    \[ \eta_{jk} > 0 \text{ iff } k \text{ is located the left or right of } j \]
  - Quality differentiation:
    \[ \eta_{jk} > 0 \text{ iff } k \text{ is located immediately above or below } j \]

- General properties of local competition models:
  - Bertrand-Nash equilibrium prices are *increasing* in the degree of differentiation (i.e. \( x_2 - x_1 \))
  - Principle of maximum differentiation: If firms endogenously choose their characteristics, they will choose to locate at \( x_1 = 0 \) and \( x_2 = 1 \).
  - In general, as the number of products goes to infinity (e.g. entry cost → 0), profit converges to zero (i.e. unlike monopolistic competition)
Hybrid model: Random-coefficients

- Most products are both horizontally and vertically differentiated.
- Example: Cars
  - Horizontal attributes: Family vs sports vs light-truck segments
  - Vertical attributes: reliability, horse-power, etc.
- The work-horse model is the random-utility model with heterogenous valuations for characteristics (aka random-coefficients):
  \[
  \max_{j \in J} x_j \beta_i - \alpha_i p_j + \epsilon_{ij}
  \]
- Note: Despite the multiplicative form, this functional form nests some Hotelling-style models.
  - Example: \( u_{ij} = \delta_j - p_j - \lambda (z_j - l_i)^2 = x_j \beta_i - \alpha p_j \).
- Let, \( A_j(x, p) = \{(\alpha_i, \beta_i, \epsilon_i)|x_j \beta_i - \alpha_i p_j + \epsilon_{ij} > x_k \beta_i - \alpha_i p_k + \epsilon_{ik}, \forall k \neq j\} \).
- Aggregate demand:
  \[
  D(p_j, p_{-j}; x) = M \times \int_{(\beta_i, \alpha_i, \epsilon_i) \in A_j(x, p)} dF(\beta_i, \alpha_i, \epsilon_i) = M \times s_j
  \]
Example: Nested-logit with demographic characteristics

- Notation:
  - $t \in (1, \ldots, T)$: Decision nodes
  - $(z_i, x_j)$: consumer/car characteristics

- Indirect utility for car $j$:

$$u_{i,j} = \sum_{t=1}^{T} x_{k(t,j)} \beta_i^t + \varepsilon_{i,k(t,j)}, \quad \text{where } \beta_i^t = \pi^t z_i.$$  

where $\varepsilon_{ik}$ is a Generalized Extreme Value (GEV) random variable, and $k(t,j)$ is the option of car $j$ in node $t$.

- At each node, the conditional probability of choosing an option takes a logit form.

- Example: 2 nodes

$$P_j = P_{j|k} P_k : \begin{cases} 
  P_j = \frac{\exp(x_j \beta_2 / \lambda_k)}{\sum_{j' \in C_2 | k} \exp(x_{j'} \beta_1 \lambda_k)} \\
  P_k = \frac{\exp(w_k \beta_1 + \lambda_k I_k)}{\sum_{l \in C_1} \exp(w_l \beta_1 I_l / \lambda_l)} \\
  I_k = \ln \sum_{j \in C_2 | k} \exp(x_j \beta_2 / \lambda_k) 
\end{cases}$$

Source: Goldberg (1995)
Example: Mixed-Logit

- The **nested-logit** model allows the (unobserved) taste for *discrete* segments to be vary across consumers.
- The **mixed-logit** model allows the (unobserved) taste for *continuous* characteristics to differ across consumers.
- Normal random-coefficient with demographics characteristics ($z_i$):

$$u_{ij} = \begin{cases} x_j \beta_i + \epsilon_{ij} & \text{If } j \neq 0 \\ \epsilon_{i0} & \text{Else.} \end{cases}$$

where $\beta_i = \pi z_i + \eta_i$, $\eta_i \sim N(0, \Sigma)$, and $\epsilon_{ij} \sim T1EV(0, 1)$.

- Aggregate market share:

$$s_j = \int \int \frac{\exp(x_j(\pi z_i + \eta_i))}{1 + \sum_{j' \in J} \exp(x_{j'}(\pi z_i + \eta_i))} dF(z_i) \phi(\eta_i | \Sigma) d\eta_i$$

- Why is it useful? This is the most flexible discrete choice model. Can approximate any random-utility model (McFadden and Train 2000).
Side note: Numerical Integration

**Challenge:** The mixed-logit model requires to numerically integrate the distribution of consumer heterogeneity.

**How?**
- Monte-Carlo integration:
  
  \[ s_j \approx \frac{1}{S} \sum_{i=1}^{S} \frac{\exp(x_j(\pi z_i + \eta_i))}{1 + \sum_{j' \in J} \exp(x_{j'}(\pi z_i + \eta_i))} \]

  where \( \{z_i, \eta_i\} \) are \( S \) random draws from \( F(z_i) \) (e.g. CPS) and \( \Phi(\eta_i|\Sigma) \).

- Refinements:
  - Halton sequence (i.e. equally-spaced percentiles)
  - Importance sampling

- Quadrature methods: Approximate (low dimensional) integrals using a finite number of points and weights
  
  \[ \int f(x)dx \approx \sum_{i=1}^{n} f(x_i)w(x_i) \]

  - References: Judd (1998), Skrainka and Judd (2011)
From McFadden to BLP

- Panel of market shares and product characteristics:

\[ \{ S_t, p_t, x_t \}_{t=1,...,T} \]

where \( t \) indexes a market, \( x_t = \{ x_{jt, 1}, \ldots, x_{jt, K} \}_{j=1,...,n_t} \) is a matrix of observed characteristics, and \( \{ p_t, s_t \} = \{ p_{jt}, s_{jt} \}_{j=1,...,n_t} \) is a matrix of endogenous prices and market shares.

- Market shares:

\[ s_{jt} = M_t \times q_{jt} \]

where \( M_t \) is the observed market size (exogenous). Examples:

- **Cars:** U.S. population
- **Cereals:** Population \( \times \) Avg. servings per month.
- **Gasoline:** Transportation needs (including car, bus, walk, etc.)
From McFadden to BLP

- Indirect utility function (Nevo 2001):

\[ u_{ijt} = \begin{cases} \sum_{k=1}^{K} \beta_{i,k} x_{jt,k} - \alpha_i p_{jt} + \xi_{jt} + \epsilon_{ij} & \text{if } j \neq 0 \\ \epsilon_{i0} & \text{else.} \end{cases} \]

\[ \beta_{i,k} = \bar{\beta}_k + z_i \pi_k + \eta_{i,k} \]
\[ \alpha_i = \bar{\alpha} + z_i \pi_p + \eta_{i,p} \]
\[ z_i \sim F(\cdot) \text{ (known)} \text{ and } \eta_i \sim N(0, \Sigma) \text{ (unknown)} \]

- **Contribution**: Berry (1994) and Berry et al. (1995) introduced a method to account for unobserved product characteristics ($\xi_{jt}$).

- **Why is it important?**
  - **Structural residual**: Allow the model to explain differences in market shares that are not accounted for by $x_{jt}$ and $F_t(\beta_i, \alpha_i)$. The alternative is to rely on measurement error (e.g. Bresnahan 1987).
  - **Simultaneity**: Recast the identification of the model into a “standard” IV problem. Previous literature relied on exogenous characteristics (e.g. Goldberg 1995), or on deterministic models (e.g. Bresnahan 1987).
  - Failing to account for unobserved product heterogeneity leads to biased estimates own and cross-elasticities.
**Inverse Demand Representation**

- **Challenge:** The error of the model enters non-linearly into the aggregate demand (outcome). Does not look like standard IV methods should work...

\[
s_{jt} = \sigma_j(\delta_t, x_t, p_t; \Sigma) = \int \int \frac{\exp(\delta_{jt} + \mu_{ij})}{1 + \sum_{j'=1}^{n_t} \exp(\delta_{j't} + \mu_{i,j'})} dF(z_i) \phi(\eta_i; \Sigma) d\eta_i
\]

where \( \delta_{jt} = x_{jt} \bar{\beta} - \bar{\alpha} p_{jt} + \xi_{jt} \) is the average value of \((j, t)\), and \( \mu_{ij} = x_{jt} \eta_{i,x} - \eta_{i,p} p_{jt} \) is the idiosyncratic value of observed characteristics.

- **Solution:** The error term is additive and separable in the inverse demand representation

\[
\xi_{jt} = \sigma_{jt}^{-1}(s_t, x_t, p_t; \Sigma) - x_{jt} \bar{\beta} + \bar{\alpha} p_{jt}, \text{ iff } \theta = \theta^0
\]

where \( \theta = (\Sigma, \bar{\beta}, \bar{\alpha}) \).

- **Important:** Berry et al. (1995) and Berry et al. (2013) show that there exists a unique solution under fairly general assumptions.

- This leads to a standard GMM estimator with IV vector \( Z_{jt} \):

\[
\min_{\theta} (Z^T \xi(\theta))^T W^{-1} (Z^T \xi(\theta))
\]
Side note: Nested-fixed point algorithm

- To estimate $\hat{\theta}$ you will need two routines: (i) a non-linear optimization algorithm to minimize the GMM objective function, and (ii) a non-linear equation solver to invert the demand system.
  - You should use standard (i.e. canned) routines for (i), but it is typically faster to code your own non-linear solver for (ii).

- Sketched of the algorithm:
  0. Guess $\Sigma^1$
  1. Invert demand: $\delta_{jt}(\Sigma^1) = \sigma_j^{-1}(s_t, p_t, x_t | \Sigma^1)$
  2. Estimate “linear parameters” using IVs $Z_{jt}$: $\delta_{jt}(\Sigma^1) = x_{jt} \bar{\beta} - \bar{\alpha} p_{jt} + \xi_{jt}$
  3. Evaluate moment conditions: $\bar{m}(\Sigma^1) = \sum_{j,t} \hat{\xi}_{jt}(\Sigma^1) Z_{jt}$
  4. Check convergence of the GMM problem. Return to [1].

- Berry et al. (1995) show that the following algorithm is a contraction mapping that can be used in step [1]:

$$\delta_{jt}^{k+1} = \delta_{jt}^k + \ln s_{jt} - \ln \sigma_j(\delta^k, p_t, x_t; \Sigma)$$

Convergence is very slow with this algorithm. It is recommended to combine this with a Newton algorithm.
What in the data identifies the model?

- Like any IV problem, we need to have enough instruments to estimate \((\bar{\beta}, \bar{\alpha}, \Sigma)\). But what is a relevant/valid instrument?

- To define the identification problem, we initially consider environments without prices.

  **Assumption:** Conditional independence of unobserved attributes and the menu of product characteristics (Berry et al. 1995)

  \[
  E(\xi_{jt} | x_t) = 0 \quad (\text{CI})
  \]

- This suggests that \(\theta\) is identified by variation in the choice-set of consumers.
  - Most identification sections describe the “ideal” experiment in which a new product enters, and steal market shares from existing products with similar attributes.

- This “red-bus/blue-bus” logic is incomplete...
The Identification Problem

- The presence of an unobservable characteristic implies that when \( x_t \) varies, so do unobservables \( \xi_t = (\xi_{1t}, \ldots, \xi_{nt}) \).

- For any hypothesis \( \theta \) we have a unique quality assignment function (Berry, Gandhi and Haile, 2013)

  \[
  s_t = \sigma(x_t, \xi_t; \theta) \iff \\
  \xi_t = \sigma^{-1}(s_t, x_t; \theta)
  \]

- Since any \( \theta \) can explain the data equally well, the model is not identified by the quality of the fit.
  
  ▶ This is what distinguish a BLP-model from a McFadden-model.

- **Identification problem:** What is the correct quality assignment model?
Identification of the Random Coefficient Model

Berry and Haile (2014)

- Partition the characteristics vector, \( x_t = \{ x_{jt}^{(1)}, x_{jt}^{(2)} \} \), such that
  \[
  u_{ijt} = \beta_1 x_{jt}^{(1)} + \xi_{jt} + \sum_{k} x_{jt,k} \eta_{ik} + \epsilon_{ijt}
  \]

- Express the average quality of products in units of \( x_{jt}^{(1)} \) to obtain a system of \textit{simultaneous equations} with \( n_t \) endogenous variables:
  \[
  x_{jt}^{(1)} = \frac{1}{\beta_1} \sigma_j^{-1} (s_t, x_t^{(2)}; \theta) + \frac{1}{\beta_1} \xi_{jt}.
  \]

- Apply CI restriction to obtain the \textit{reduced-form} of the model:
  \[
  x_{jt}^{(1)} = E \left[ \frac{1}{\beta_1} \sigma_j^{-1} (s_t, x_t^{(2)}; \theta) \mid x_t \right] + 0 = g_j(x_t; \theta)
  \]

- If \( x_t \) (\( n_t \cdot K \) variables) is complete for \( \{ s_t, x_t^{(2)} \} \) (\( n_t \cdot K \) variables) then \( F(\eta) \) is non-parametrically identified.
  - Completeness is the \textit{non-parametric} analogue of the rank condition for parametric models.
Special Case: Nested-Logit

- **Indirect utility:**

\[ u_{jt} = x_{jt}\beta_x + \sum_{k} \eta_{ik}1(x_{jt,k} = 1) + \xi_{jt} + \varepsilon_{ijt} \]

**Note:** Nested-logit is equivalent to a model with random-coefficients on product segments dummies.

- Linear quality assignment regression (Berry, 1994):

\[ \sigma^{-1}_j(s_t,x_t;\theta^0) = \ln(s_{jt}/s_{0t}) - \lambda \ln(s_{jt}/s_{g(j),t}) - x_{jt}\beta = \xi_{jt} \]

\[ \leftrightarrow \ln(s_{jt}/s_{0t}) = x_{jt}\beta_x + \lambda \ln(s_{jt}/s_{g(j),t}) + \xi_{jt} \]

- **Relevant and Valid IVs** are the characteristics of competitors within the same nest (e.g. Verboven (1996), Bresnahan et al. (1997))
Simultaneous Equation for the Non-Linear Case

- To define relevant instruments, we can linearize the system around the true \( \theta^0 \) (Jorgenson and Laffont (1974) and Amemiya (1974)):

\[
\sigma_{jt}^{-1}(s_t, x_t; \theta) = \sum_k (\Sigma_k - \Sigma^0_k) \frac{\partial \sigma_{jt}^{-1}(s_t, x_t; \theta^0)}{\partial \theta_k} + \xi_{jt} + \text{Error}
\]

\[
= J_{jt}(s_t, x_t; \theta^0)b + u_{jt}
\]

This leads to a linear IV regression: Gauss-Newton Regression.

**Important:** \( ||\hat{b}^{IV}|| = 0 \) defines any GMM estimator of \( \theta \).

- Therefore, the **most relevant** IV for \( \theta_k \) involves calculating the best predictor \( J_{jt} \) given \( x_t \) (Amemiya (1977), Chamberlain (1987)):

\[
h_{j,k}(x_t; \theta^0) = E \left[ J_{jt,k}(s_t, x_t; \theta^0) | x_t \right]
\]
The linear representation can be used to construct a Gauss-Newton algorithm to estimate \({\hat{\theta}}\).

\[
\sigma_{jt}^{-1}(s_t, x_t; \theta) = \sum_k (\theta_k - \theta_0^k) \frac{\partial \sigma_j^{-1}(s_t, x_t; \theta^0)}{\partial \theta_k} + \xi_{jt} + \text{Error}
\]

\[
= J_{jt}(s_t, x_t; \theta^0)b + u_{jt}
\]

Iteration \(k \geq 1:\)

1. Invert demand: \(\sigma_{jt}^{-1}(s_t, x_t|\theta^{k-1})\) and \(J_{jt}(s_t, x_t|\theta^{k-1})\)
2. Estimate \({\hat{b}^k}\) by 2SLS.
3. If \(|\hat{b}^k| < \varepsilon\) stop. Else, update \(\theta^k = \theta^{k-1} + \hat{b}^k\), and repeat steps (1)-(3).

With strong instruments, this procedure typically requires less than 5 iterations.

- Weak IVs lead to severe numerical problems (e.g. Knittle and Metaxoglou (2014), Dube et al. (2012))
- Why? The central-limit theorem does no hold, and the quadratic approximation is bad.
Optimal IV

- The “optimal” IV for $\Sigma$ is (Amemyia (1977), Chamberlain (1987)):

$$D_{j,k} (x_t; \Sigma^0) = E \left[ \frac{\partial \sigma^{-1}_j (s_t, x_t^{(2)}; \Sigma^0)}{\partial \Sigma_k} \bigg| x_t \right]$$

- **How?** Approximate the optimal IV by estimating a non-parametric first-stage regression (Newey, 1990)

- **Curse of Dimensionality:** $D_{j,k} (x_t)$ is a $j$-specific function of the entire menu of characteristics.
  - First order: $D_{j,k}(x_t; \Sigma^0) \approx \sum_{j=1}^{n_t} x_{jt} a_j$
  - Impossible: BLP (1995)

- **Bottom line:** If we can’t estimate the first-stage, how can we construct powerful instruments?
What does the characteristic structure imply for the reduced-form of the model?

- Market-structure facing product \( j \) (dropping \( t \)):

\[
(w_j, w_{-j}) \equiv \left( (\delta_j, x_j^{(2)}), (\delta_{-j}, x_{-j}^{(2)}) \right)
\]

- Properties of the linear-in-characteristics model:

  - Symmetry:

\[
\sigma_j (w_j, w_{-j}) = \sigma_k (w_j, w_{-j}) \quad \forall k \neq j
\]

  - Anonymity:

\[
\sigma (w_j, w_{-j}) = \sigma (w_j, w_{\rho(-j)}) \quad \forall \rho
\]

  - Translation invariant: for any \( c \in \mathbb{R}^K \)

\[
\sigma (w_j + (0, c), w_{-j} + (0, c)) = \sigma (w_j, w_{-j})
\]
Re-Express the Demand System

- Express the “state” of the market in differences relative to $j$ and treat the outside option just like any other product.
  - Characteristic differences:
    $$d_{j,k}^{(2)} = x_k^{(2)} - x_j^{(2)}$$
  - New normalization:
    $$\tau_j = \frac{\exp(\delta_j)}{1 + \sum_{j'} \exp(\delta_{j'})}, \forall j = 0, \ldots, n.$$  
  - Product $k$ attributes: $\omega_{j,k} = (\tau_k, d_{jt,k}^{(2)})$

- Demand for product $j$ is a fully exchangeable function of $\omega_j$:
  $$\sigma (w_j, w_{-j}) = D(\omega_j)$$

where $\omega_j = \{\omega_{j,0}, \ldots, \omega_{j,j-1}, \omega_{j,j+1}, \ldots, \omega_{j,n}\}$. 

Demand for Differentiated Products

Choice of Instruments
Main Theory Result

- Define the *exogenous* state of the market facing product $j$:

$$
d_{j,k} = x_k - x_j$$

$$d_j = (d_{j,0}, \ldots, d_{j,j-1}, d_{j,j+1}, \ldots, d_{j,n})$$

### Theorem

If the distribution of $\{\xi_j\}_{j=1,\ldots,n}$ is exchangeable, then the reduced form becomes

$$D_{j,k} (x; \Sigma^0) = E \left[ \frac{\partial \sigma^{-1}_j (s, x^{(2)}; \Sigma^0)}{\partial \Sigma_k} \bigg| x \right] = h_k (d_j)$$

where $h_k$ is a **symmetric** function of the state vector.

- **Implication**: $h_k$ is a *vector symmetric function* (see Briand 2009)
Why is it useful?

1. **Curse of dimensionality:** In any fixed order approximation the number of basis functions is *independent* of the number of products.

2. **Examples:** Basis functions
   - First-order (i.e. market):
     \[ h(d_j) \approx \sum_{k \neq j} a_k d_{jt,k} = a \times \left[ \sum_{k \neq j} d_{jt,k} \right] \left( \equiv \sum_{j' \neq j} x_{jt} \right) \]
   - Second-order (i.e. distance):
     \[ \sum_{j' \neq j} \left( d_{jt,j'}^k \right)^2, \quad \forall k = 1, \ldots, K \]
   - Histogram basis (i.e. nested-logit or hotelling):
     \[ \sum_{j' \neq j} 1(|d_{jt,j'}^k| < \kappa_k) d_{jt,j'}, \quad \forall k = 1, \ldots, K \]
Let $A_j(x_t)$ be an $L$ vector of basis functions summarizing the empirical distribution of characteristic differences: $\{d_{jt,k}\}_{k=0,...,n_t}$.

**Differentiation IV:** These functions are moments describing the relative isolation of each product in characteristic space.

**Donald, Imbens, and Newey (2003):** Using basis functions directly as IVs, is asymptotically equivalent to approximating the optimal IV.

- Recommended practice is to use low-order basis functions (Donald, Imbens, and Newey 2008).

Alternatively use Lasso methods with higher order polynomials (e.g. Belloni, Chernozhukov, and Hansen (2012)).
Side Note: Demographic Panel

- In many settings, product characteristics are fixed across markets, but the distribution of consumer types vary (e.g. Nevo 2001).
- Focus on a single characteristics $x_j$
- **Assumption:** Consumer heterogeneity $\eta_{it}$ follows distribution of a demographic can be decomposed as

$$\eta_{it} = \mu_t + \sigma_t \nu_i$$

where $(\mu_t, \sigma_t)$ are known, and $\Pr(\nu_i < x) = F(x)$ is common across $t$.
- Preferences

$$u_{ijt} = \delta_{jt} + \theta \eta_{it} x_j + \varepsilon_{ijt}$$

$$= \tilde{\delta}_{jt} + \theta \nu_i \tilde{x}_{jt} + \varepsilon_{ijt}$$

where $\tilde{x}_{jt} = \sigma_t x_j$, and $\delta_{jt} = x_j (\beta + \theta \mu_t) + \xi_{jt} = x_j \beta_t + \xi_{jt}$.
- Differentiation IVs can be constructed as before: $d_{jt,k}^x = \tilde{x}_{kt} - \tilde{x}_{jt}$.
Illustration: A Motivating Example

- One dimension random coefficient model:
  \[ u_{ijt} = \beta_0 + \beta_1 x_{jt}^{(1)} + \beta_2 x_{jt}^{(2)} + \xi_{jt} + \sigma x_i x_{jt}^{(2)} + \varepsilon_{ijt}. \]

- Unobserved heterogeneity:
  - \( v_i \sim N(0, 1) \)
  - Numerical integration: 10 points Gauss–Hermite quadrature.

- Data:
  - Panel structure: 100 markets \( \times \) 15 products
  - Characteristics: \( (\xi_{jt}, x_{jt}) \sim N(0, I) \).
  - Dimension: \( |x_{jt}| = 2 \)
Identification in a Picture

Market IV: Sum of rivals’ characteristics

(A) Residual quality at $\sigma_x = \sigma_0^x$

(B) Residual quality at $\sigma_x = 0$

- **Moment restriction:** Independence of $\xi_{jt}$ and the sum of rivals characteristics cannot distinguish (A) from (B)

- **Weak identification:** The moment restrictions are “almost” satisfied at $\sigma_x \neq \sigma_0^x$ (Stock and Wright, 2000).
Identification in a Picture

Differentiation IV: Number of competitors located within 1 std-dev of $x_{jt}^{(2)}$

(A) Residual quality at $\sigma_x = \sigma_0^x$

![Graph A](image)

Note: Regression $R^2 = 0$. 

(B) Residual quality at $\sigma_x = 0$

![Graph B](image)

Note: Regression $R^2 = .11$.

- **Moment restriction:** Independence of $\xi_{jt}$ and the *number of close competitors* rules out (B).
Monte-Carlo Experiment

- Random coefficient model:

\[ u_{ijt} = \delta_{jt} + \sum_{k=1}^{K} v_{ik} x_{jt,k}^{(2)} + \varepsilon_{ijt}, \quad v_i \sim N(0, \sigma_x^2 I). \]

- Data:
  - Panel structure: 100 markets $\times$ 15 products
  - Characteristics: $(\xi_{jt}, x_{jt}) \sim N(0, I)$.
  - Dimension: $|x_{jt}| = K + 1$
  - Monte-Carlo replications $= 1,000$

- Instruments $(K + 1)$:
  - Market IVs (first-order): $A_j(x_t) = \sum_{l \neq j}^{n_t} x_{lt,k}, \forall k = 1, \ldots, K$
  - Diff. IVs (2nd-order): $A_j(x_t) = \sum_{j'=1}^{n_t} (d_{jt,j'}^k)^2, \forall k = 1, \ldots, K$
Result 1: Weak IV Problem

Specification: One dimension of heterogeneity
### Result 2: Weak IVs with Multiple Dimensions

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>Dimensions of consumer heterogeneity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>$\sigma_1$</td>
<td>2.012</td>
</tr>
<tr>
<td></td>
<td>(1.3)</td>
</tr>
<tr>
<td>$\sigma_2$</td>
<td>1.978</td>
</tr>
<tr>
<td></td>
<td>(1.302)</td>
</tr>
<tr>
<td>$\sigma_3$</td>
<td>1.992</td>
</tr>
<tr>
<td></td>
<td>(1.353)</td>
</tr>
<tr>
<td>$\sigma_4$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Nb. IVs  | 2 | 3 | 4 | 5 |

Parenthesis: RMSE. Starting at true parameter = 2.
### Result 3: Differentiation IVs with Multiple Dimensions

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>Dimensions of consumer heterogeneity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>$\sigma_1$</td>
<td>2.002</td>
</tr>
<tr>
<td></td>
<td>(.107)</td>
</tr>
<tr>
<td>$\sigma_2$</td>
<td>1.995</td>
</tr>
<tr>
<td></td>
<td>(.108)</td>
</tr>
<tr>
<td>$\sigma_3$</td>
<td>1.995</td>
</tr>
<tr>
<td></td>
<td>(.109)</td>
</tr>
<tr>
<td>$\sigma_4$</td>
<td>1.995</td>
</tr>
<tr>
<td></td>
<td>(.11)</td>
</tr>
<tr>
<td>$\sigma_5$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma_6$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Nb. IVs</td>
<td>2</td>
</tr>
</tbody>
</table>

Parenthesis: RMSE. Starting at true parameter = 2.
## Result 4: Weak IVs and Numerical Problems

Simulation: 1,000 samples and 10 random starting values each

<table>
<thead>
<tr>
<th>Differentiation IV</th>
<th>Market IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Est. RMSE</td>
<td>Est. RMSE</td>
</tr>
</tbody>
</table>

| \( \sigma_1 \) | 1.994 | 0.111 | 2.107 | 1.536 |
| \( \sigma_2 \) | 1.992 | 0.111 | 2.053 | 1.477 |
| \( \sigma_3 \) | 1.994 | 0.113 | 2.088 | 1.618 |
| \( \sigma_4 \) | 1.995 | 0.110 | 2.131 | 1.640 |

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Simplex</th>
<th>Simplex</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nb. IVs</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>CPU time (sec)</td>
<td>0.718</td>
<td>33</td>
</tr>
<tr>
<td>Local optima (fraction)</td>
<td>0.002</td>
<td>0.53</td>
</tr>
<tr>
<td>Global min/Local min</td>
<td>46.6</td>
<td>53</td>
</tr>
</tbody>
</table>
**Result 4: Weak IVs and Numerical Problems**

Simulation: 1,000 samples and 10 random starting values each

<table>
<thead>
<tr>
<th>σ</th>
<th>Differentiation IV Est.</th>
<th>RMSE</th>
<th>Market IV Est.</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>σ₁</td>
<td>1.994</td>
<td>0.111</td>
<td>2.107</td>
<td>1.536</td>
</tr>
<tr>
<td>σ₂</td>
<td>1.992</td>
<td>0.111</td>
<td>2.053</td>
<td>1.477</td>
</tr>
<tr>
<td>σ₃</td>
<td>1.994</td>
<td>0.113</td>
<td>2.088</td>
<td>1.618</td>
</tr>
<tr>
<td>σ₄</td>
<td>1.995</td>
<td>0.110</td>
<td>2.131</td>
<td>1.640</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Simplex</th>
<th>Simplex</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nb. IVs</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>CPU time (sec)</td>
<td>0.718</td>
<td>33</td>
</tr>
<tr>
<td>Local optima (fraction)</td>
<td>0.002</td>
<td>0.53</td>
</tr>
<tr>
<td>Global min/Local min</td>
<td>46.6</td>
<td>53</td>
</tr>
</tbody>
</table>
Experiment 2: Correlated Random Coefficients

- Consumer heterogeneity:

\[ \eta_i \sim N(0, \Sigma) \]

Note: 4 dimensions \( \Rightarrow \) 10 non-linear parameters (choleski)

- Panel structure:

100 markets \( \times \) 50 products

- **Differentiation IVs:** Second-order polynomials (with interactions):

\[
\sum_{j'=1}^{n_t} \left( d_{jt,j'}^k \right)^2 \quad \text{and} \quad \sum_{j'=1}^{n_t} \left( d_{jt,j'}^k \times d_{jt,j'}^l \right)
\]

for all characteristics \( k \) and \( l \).
Simulation Results: Unrestricted Covariances

<table>
<thead>
<tr>
<th></th>
<th>( \Sigma_1 )</th>
<th>( \Sigma_2 )</th>
<th>( \Sigma_3 )</th>
<th>( \Sigma_4 )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Estimated parameters:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \Sigma_1 )</td>
<td>4.066</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.214)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \Sigma_2 )</td>
<td>-2.080</td>
<td>4.268</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.132)</td>
<td>(.242)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \Sigma_3 )</td>
<td>1.993</td>
<td>0.697</td>
<td>4.502</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.143)</td>
<td>(.154)</td>
<td>(.212)</td>
<td></td>
</tr>
<tr>
<td>( \Sigma_4 )</td>
<td>-1.147</td>
<td>-0.405</td>
<td>-0.458</td>
<td>3.050</td>
</tr>
<tr>
<td></td>
<td>(.123)</td>
<td>(.124)</td>
<td>(.135)</td>
<td>(.204)</td>
</tr>
<tr>
<td><strong>True parameters:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \Sigma_1 )</td>
<td>4.064</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \Sigma_2 )</td>
<td>-2.083</td>
<td>4.270</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \Sigma_3 )</td>
<td>1.995</td>
<td>0.689</td>
<td>4.505</td>
<td></td>
</tr>
<tr>
<td>( \Sigma_4 )</td>
<td>-1.152</td>
<td>-0.398</td>
<td>-0.462</td>
<td>3.071</td>
</tr>
<tr>
<td><strong>CPU time (sec)</strong></td>
<td>3.151</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Parenthesis: RMSE. Starting values = true. Algorithm: Gauss-Newton.
Endogenous Prices

- Berry (1994) discuss the simultaneity of prices and unobserved attributes ($\xi_{jt}$). Conditional on $\Sigma$, this is a “simpler” linear-IV problem:

$$\delta_{jt}(\Sigma) = x_{jt}\bar{\beta} - \bar{\alpha}p_{jt} + \xi_{jt}$$

**Note:** Since the inverse demand is linear, we can easily include fixed-effects in the model to control for time invariant unobservables.

- The literature has proposed several alternatives:
  - **BLP-IVs:** Characteristics of products owned by the same firm. Example:
    $$w_{jt} = \sum_{j' \in F_j} x_{jt}$$
  - **Hausman-IVs:** Price of same and competing brands in neighboring geographic markets. Why? If the demand shock $\xi_{jt}$ is independent across markets, but production plants deliver to several markets, prices will be spatially correlated (i.e. cost shocks). E.g.: Nevo (2001).
  - **Cost IVs:** Input prices, shipping distance, merger (e.g. Miller and Weinberg (2016))
Supply Restrictions: Pricing game and cost function

- Equilibrium pricing restrictions can be very informative about demand parameters.
- Assume the following constant marginal cost expression:

\[ \ln(mc_{jt}) = w_{jt} \gamma + \omega_{jt} \]

- The multi-product Bertrand-Nash equilibrium price vector is characterized by the following FOC for all \( j \).

\[ s_{jt} + \sum_{k \in F_{jt}} (p_{kt} - mc_{kt}) \frac{\partial \sigma_{k}(\delta_{t}, p_{t}, x_{t}; \Sigma)}{\partial p_{jt}} = 0 \]

where \( F_{jt} \) is the set of products owned by the same firm as \( j \)’s.
Supply Restrictions: Pricing game and cost function

- Matrix representation:

\[ s_t + (\Delta_t \ast \Omega_t)(p_t - mc_t) = 0 \]

where \( \Delta_{jk,t} = \frac{\partial \sigma_k(\delta_t, p_t, x_t; \Sigma)}{\partial p_{jt}} \), and \( \Omega_{jk,t} = 1 \) if \( (jk) \) are owned by the same company (zero otherwise).

- Given \( \Sigma \), the marginal cost vector can be recovered from the data:

\[ mc_t(\Sigma, \gamma) = w_t \gamma + \omega_t = p_t + (\Omega_t \ast \Delta_t)^{-1} \sigma(\delta_t, p_t, x_t|\Sigma) \]

- We can therefore form a joint vector of moment conditions:

\[ E \left[ \begin{array}{c} \xi_{jt}(\theta)z_{jt} \\ \omega_{jt}(\theta, \gamma)z_{jt} \end{array} \right] = 0 \]

- The demand parameters are therefore identified by the variation coming from the IVs, and the cross-equation restrictions implied by the pricing model.
Illustration: Vertical Differentiation

- **Demand:** Mixed-logit with vertical differentiation
  - Indirect utility
    \[ u_{ijt} = \delta_{jt} - \alpha_i p_{jt} + \epsilon_{ijt} \]
    where \( \alpha_i = \sigma_p y_i^{-1} \), and \( \log(y_i) \sim N(\mu_y, \sigma_y^2) \) (known).

- **Supply:** Bertrand-Nash with multi-product firms
  - Constant marginal cost: \( mc_{jt} = \gamma_0 + \gamma_x x_{jt} + \omega_{jt} \)
How to incorporate endogenous prices?

- **Price instruments:** \( w_t = \{w_{jt}\}_{j=1,...,n_t} \)

- **Reduced-form of the model:**

\[
D_j(x_t, p_t, w_t; \sigma_p^0) = E \left( \frac{\partial \sigma_j^{-1}(s_t, x_t, p_t | \sigma_p^0)}{\partial \sigma_p} \Big| x_t, w_t \right) \neq h(d_j^x, d_j^p)
\]

- **Solution:** Use insights from Berry et al. (1999) and Reynaert & Verboven (2013) to take expectation for price inside:

\[
D_j(x_t, p_t, w_t; \sigma_p^0) \approx D_j(x_t, \hat{p}_t, w_t; \sigma_p^0) = h\left(d_j^x, \hat{d}_j^p\right)
\]

I.e., \( d_{jt,k}^p = p_{kt} - p_{jt} \) is replaced by \( \hat{d}_{jt,k}^p = E(p_{kt} | w_{kt}) - E(p_{jt} | w_{jt}) \).
BLP (1995) Instruments

- BLP’s original basis function:

\[ Z_{jt} = \left\{ w_{jt}, \sum_{j' \in F_{ft}} w_{j',t}, \sum_{j' \notin F_{ft}} w_{j',t} \right\} \]

where \( w_{jt} = (1, x_{jt}, \omega_{jt}) \), and \( F_{ft} \) is the set of products controlled by firm \( f \).

- This corresponds to the first moment of characteristic differences

\[ Z'_{jt} = \left\{ w_{jt}, \sum_{j' \in F_{ft}} d^{w}_{j,t,k}, \sum_{j' \notin F_{ft}} d^{w}_{j',t} \right\} \]

where \( d^{w}_{jt,k} = \left( 1, x_{jt,k}, \omega_{jt,k} \right) \), and \( \text{span}(Z) = \text{span}(Z') \).
Differentiation IVs

- **Characteristic differences:** Let $d_{jt,k} = \hat{p}_{kt} - \hat{p}_{jt}$ denote the *exogenous* price differences between $j$ and $k$, where:

  $$\hat{p}_{jt} = \hat{\pi}_0 + \hat{\pi}_1 x_{jt} + \hat{\pi}_2 \omega_{jt}.$$ 

- **Two basis functions:**
  1. *Second moment:* Euclidian distance

     $$z_{jt} = \left\{ w_{jt}, \sum_{j' \in F_t} d_{j',jt}^2, \sum_{j' \notin F_t} d_{j',jt}^2 \right\}$$

  2. *Histogram:* Characteristics of local competitors

     $$z_{jt} = \left\{ w_{jt}, \sum_{j' \in F_t, |d_{j',jt}| < \text{sd}(d)} w_{j't}, \sum_{j' \notin F_t, |d_{j',jt}| < \text{sd}(d)} w_{j't} \right\}$$
Weak IV problem is very severe without cost-shifters

Panel structure: $\bar{J} = 50$ and $T = 10$

Note: True value = 20
Simulation Results

<table>
<thead>
<tr>
<th></th>
<th>IV: Sum of charact.</th>
<th></th>
<th>IV: Local competitors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>w/o cost w/ cost</td>
<td>w/o cost w/ cost</td>
<td></td>
</tr>
<tr>
<td>$\beta_p$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>0.46 1.08</td>
<td>1.00</td>
<td>1.02</td>
</tr>
<tr>
<td>RMSE</td>
<td>2.19 1.32</td>
<td>0.22</td>
<td>0.18</td>
</tr>
<tr>
<td>$\sigma_p$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>13.24 17.47</td>
<td>19.10</td>
<td>19.68</td>
</tr>
<tr>
<td>RMSE</td>
<td>10.84 7.95</td>
<td>3.93</td>
<td>1.51</td>
</tr>
</tbody>
</table>
The weak IV problem does not go away in large samples...

Sample size: Solid = 500, Long dash = 1,000, Dash = 2,500.
Natural Experiments

- **Hotelling example:**

\[
u_{im} = \xi_{jm} - \lambda(\eta_i - x_{jm})^2 + \epsilon_{ijm}\]

and \(E(\xi_{jt}|x_{jt}) \neq 0\) (i.e. endogenous locations).

- **Natural experiment:** Exogenous entry of a new product \((x' = 5)\)

- **Three-way panel:**
  - Product \(j\), market \(m\), and time \((t = 0, 1)\).
  - Treatment variable:

\[
D_{jm} = 1(|x_{jm} - 5| < \text{Cutoff})
\]

- “Ideal” first-stage: Difference-in-difference regression

\[
\frac{\partial \sigma_j^{-1}(s_t, x_t, p_t|\theta^0)}{\partial \theta} = \hat{\mu}_{jm} + \hat{\tau}_t + \hat{\gamma}D_{jm} \times 1(t = 1) + \hat{e}_{jmt}
\]
Natural Experiment: Hotelling Example

DGP: $\delta_{jmt} = \bar{\xi}_{jm} + \Delta \xi_{jmt}$, where $E(\bar{\xi}_{jm}|x_m) \neq 0$

- “Diff-in-Diff” specification:

$$z_{jmt} = \{\text{Product Dummy}_{jm}, 1(t = 1), 1(|x_{jm} - 5| < 1)1(t = 1)\}$$
Evaluating the Relevance of Instruments

**Ex-post:** First-stage test

- First-stage of the non-linear GMM problem at $\hat{\theta}^{gmm}$:

$$
\frac{\partial \sigma_{jt}^{-1}(s_t, x_t, p_t | \hat{\theta}^{gmm})}{\partial \theta} = \pi_x x_{jt} + \pi_z z_{jt} + e_{jt}
$$

- Standard tests for weak instruments in linear models can be used to test the relevance of $z_{jt}$ (e.g. IVREG2 in STATA).

**Ex-ante:** IIA test

- A strong instrument for $\Sigma$ is able to reject the wrong model (Stock and Wright, 2000)
- Under $H_0 : \Sigma = 0$, the inverse demand equation is independent of $x_{-j}$:

$$
\sigma_{jt}^{-1}(s_t, x_t; \Sigma = 0) = \ln s_{jt}/s_{0t} = x_{jt}\beta + z_{jt}\gamma + \xi_{jt}
$$

- Standard test statistics for $H_0 : \gamma = 0$, can be used to test null hypothesis of IIA preferences
- With endogenous prices, this test is equivalent to the J-test at $\Sigma = 0$. 

Demand for Differentiated Products

Testing for Weak IVs
Motivation:
- Heterogeneity parameters are often imprecisely estimated with aggregate data solely
- Propose to add “micro-moment” conditions from consumer surveys
- Originally proposed by Petrin (2002)

Additional moment: Difference between predicted and observed conditional market share of “large” cars for large/small families.

Why is it not redundant? We’re inverting the aggregate market shares, not the conditional market shares.
Empirical Specification

- Indirect utility combines demographics and normal random-coefficients:

\[ u_{ij} = \alpha_i \ln(y_i - p_j) + x_j \beta + \sum_k \gamma_k \nu_{ik} x_{ik} + \xi_j + \epsilon_{ij} \]

- Where,
  - Price sensitivity \( \alpha_i \) varies across income groups
  - Family size categories:

\[ \gamma_{ic} = \gamma_c \ln(\text{Family size}_i) \nu_i \]

where \( c \) indicates minivan, station wagons, full-size vans, and SUVs.
Moment Conditions

- IV moments: Combined demand and supply restrictions

\[
\frac{1}{n} \sum_{j,t} \xi_{jt} Z_{jt} = 0
\]

\[
\frac{1}{n} \sum_{j,t} \omega_{jt} Z_{jt} = 0
\]

- Micro-moments:
  - Share of new vehicles purchased by income group (3)
  - Average family size per car type
  - Average age conditional on buying each car type.

- All three sets of moments are “stacked” to form vector \( G_n(\theta, Z) \).
### TABLE 2

**Average Consumer Characteristics for the United States and Selected Subpopulations, 1987–92**

<table>
<thead>
<tr>
<th>United States</th>
<th>Purchasers of</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>New Vehicles</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td><strong>36,113</strong></td>
</tr>
<tr>
<td><strong>Standard Deviation</strong></td>
<td><strong>21,255</strong></td>
</tr>
<tr>
<td><strong>Income</strong></td>
<td><strong>23,728</strong></td>
</tr>
<tr>
<td><strong>Family size</strong></td>
<td><strong>1.53</strong></td>
</tr>
<tr>
<td><strong>Midage</strong></td>
<td><strong>.49</strong></td>
</tr>
</tbody>
</table>

**Source.** — Consumer Expenditure Survey.

**Note.** — Income is measured in 1982–84 CPI-adjusted dollars. Family size is the number of household members. Midage is a binary variable for the age of the head of household between 30 and 60 inclusive.
<table>
<thead>
<tr>
<th>Variable</th>
<th>OLS Logit (1)</th>
<th>Instrumental Variable Logit (2)</th>
<th>Random Coefficients (3)</th>
<th>Random Coefficients and Microdata (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. Price Coefficients (α’s)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \alpha_1 )</td>
<td>.07**</td>
<td>.13**</td>
<td>4.92**</td>
<td>7.52**</td>
</tr>
<tr>
<td></td>
<td>(.01)**</td>
<td>(.01)**</td>
<td>(9.78)**</td>
<td>(1.24)**</td>
</tr>
<tr>
<td>( \alpha_2 )</td>
<td>11.89</td>
<td>31.13</td>
<td>(21.41)**</td>
<td>(4.07)**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \alpha_3 )</td>
<td>37.92</td>
<td>34.49</td>
<td>(18.64)**</td>
<td>(2.56)**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B. Base Coefficients (β’s)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-10.03</td>
<td>-10.04</td>
<td>-12.74**</td>
<td>-15.67**</td>
</tr>
<tr>
<td></td>
<td>(.32)**</td>
<td>(.34)**</td>
<td>(5.65)**</td>
<td>(4.39)**</td>
</tr>
<tr>
<td>Horsepower/weight</td>
<td>1.48</td>
<td>3.78</td>
<td>3.40</td>
<td>-2.83</td>
</tr>
<tr>
<td></td>
<td>(.34)**</td>
<td>(.44)**</td>
<td>(39.79)**</td>
<td>(8.16)</td>
</tr>
<tr>
<td>Size</td>
<td>3.17</td>
<td>3.25</td>
<td>4.60</td>
<td>4.80</td>
</tr>
<tr>
<td></td>
<td>(.26)**</td>
<td>(.27)**</td>
<td>(24.64)**</td>
<td>(3.57)**</td>
</tr>
<tr>
<td>Air conditioning standard</td>
<td>-.20</td>
<td>.21</td>
<td>-1.97</td>
<td>3.88</td>
</tr>
<tr>
<td></td>
<td>(.06)**</td>
<td>(.08)**</td>
<td>(2.23)**</td>
<td>(2.21)**</td>
</tr>
<tr>
<td>Miles/dollar</td>
<td>.18</td>
<td>.05</td>
<td>-.54</td>
<td>-15.79</td>
</tr>
<tr>
<td></td>
<td>(.06)**</td>
<td>(.07)**</td>
<td>(3.40)**</td>
<td>(.87)**</td>
</tr>
<tr>
<td>Front wheel drive</td>
<td>.32</td>
<td>.15</td>
<td>-5.24</td>
<td>-12.32</td>
</tr>
<tr>
<td></td>
<td>(.05)**</td>
<td>(.06)**</td>
<td>(3.09)**</td>
<td>(2.36)**</td>
</tr>
<tr>
<td>Minivan</td>
<td>.09</td>
<td>-.10</td>
<td>-4.34</td>
<td>-5.65</td>
</tr>
<tr>
<td></td>
<td>(.14)</td>
<td>(.15)</td>
<td>(13.16)</td>
<td>(.68)**</td>
</tr>
<tr>
<td>Station wagon</td>
<td>-1.12</td>
<td>-1.12</td>
<td>-20.52</td>
<td>-1.31</td>
</tr>
<tr>
<td></td>
<td>(.06)**</td>
<td>(.07)**</td>
<td>(36.17)</td>
<td>(.36)**</td>
</tr>
<tr>
<td>Sport-utility</td>
<td>-.41</td>
<td>-.61</td>
<td>-3.10</td>
<td>-4.38</td>
</tr>
<tr>
<td></td>
<td>(.09)**</td>
<td>(.10)**</td>
<td>(10.76)</td>
<td>(.41)**</td>
</tr>
<tr>
<td>Full-size van</td>
<td>-1.73</td>
<td>-1.89</td>
<td>-28.54</td>
<td>-5.26</td>
</tr>
<tr>
<td>% change GNP</td>
<td>.03</td>
<td>.03</td>
<td>.08</td>
<td>.24</td>
</tr>
<tr>
<td></td>
<td>(.01)**</td>
<td>(.01)**</td>
<td>(.02)**</td>
<td>(.02)**</td>
</tr>
</tbody>
</table>

**NOTE.** Standard errors are in parentheses. A quadratic time trend is included in all specifications.

* Zstatistic >1.
** Zstatistic >2.
### TABLE 8

**Average Compensating Variation Conditional on Minivan Purchase, 1984:**

1982–84 CPI-Adjusted Dollars

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

**Note.**—Compensating variation is evaluated at equilibrium prices without minivans. Decomposition of compensation is the average difference in the value of observed and unobserved characteristics between first and second choices. For logit models, the purchase decision is independent of income, so mean purchaser income is mean U.S. household income.


Quantifying the benefits of new products: The case of the minivan.
*Journal of Political Economy* 110, 705.

Mergers, brand competition and the price of a pint.

Spatial price competition: A semiparametric approach.
*Econometrica* 70(3), 1111–1153.

Relaxing price competition through product differentiation.

working paper CWP03, CENMAP.

*Discrete Choice Methods with Simulation.*
Cambridge University Press.