Demand for Differentiated Products: Housing, Hedonic and Product Complementarity

Jean-François Houde
Cornell University & NBER

September 26, 2016

Sorting and Hedonic Models

- The goal of both approaches' is to estimate the average *willingness-to-pay* for product attributes
- **Hedonic model:**
  - **Assumption:** All possible combinations of product characteristics are available
  - Hedonic choice model: Let $X_i = \{x_{i1}, x_{i2}, \ldots, x_{iK}\}$ be a bundle of characteristics, and $p(X_i)$ be the **hedonic** price function (equilibrium).
  - Consumers solve the following problem:

$$\max_{X_i} U(X_i) + y_i - p(X_i)$$

  - Optimality of X_i implies that for each attribute k, the following conditions are satisfied separately for each attribute:

$$\frac{\partial U(X_i)}{\partial x_{ik}} = \frac{\partial p(X_i)}{\partial x_{ik}}, \forall k$$

$$U(x_{ik}, x_{i,-k}) - p(x_{ik}, x_{i,-k}) \geq U(x'_{ik}, x_{i,-k}) - p(x'_{ik}, x_{i,-k}), \forall k$$

Sorting and Hedonic Models

- **Sorting/discrete choice model:**
  - Consumers choose among a finite set of bundles that might not cover all combinations possible:

$$\max_{j \in J} X_{ij}\beta_i - \alpha_i p_j + \xi_j + \varepsilon_{ij}$$

  - Note that adding ε_{ij} makes each product "unique", and therefore inconsistent with the hedonic assumption.
  - When ϵ_{ij} is not included, the randomness in choices across consumers arises only because of heterogeneous taste of attributes. This is the **pure characteristic** model.

$$i \to j \text{ iff } X_{ij}\beta_i - \alpha_i p_j + \xi_j \geq X_{ik}\beta_i - \alpha_i p_k + \xi_k \forall k \neq j$$

- In both frameworks, it is important to account for heterogenous taste **and** omitted attributes (ξ_{jt})

Sorting and Demand for School Quality

Reference: Bayer, Ferreira, and McMillan (2007)

- There exists a large literature on measuring the value school quality, based on hedonic house price regression:

$$p_h = \alpha \text{Test score}_h + X_h\beta + e_h$$

  - $\hat{\alpha} =$ WTP for quality? Good schools are typically located in nice neighborhoods with rich (and white) people, nice houses, park, family-oriented amenities, etc.
- Sandra Black (QJE, 1999): Unobserved attributes explain most of the correlation between house prices and school quality.
- Identification strategy: Boundary discontinuity design (BDD):

$$p_h = \alpha \text{Test score}_h + X_h\beta + \mu_{b(h)} + u_h$$

  where $\mu_{b(h)}$ is a school boundary fixed-effect.
- **Questions:** If households sort based on heterogenous tastes, is the hedonic coefficient measuring the average WTP for school quality?

Test scores and house prices are clearly discontinuous...

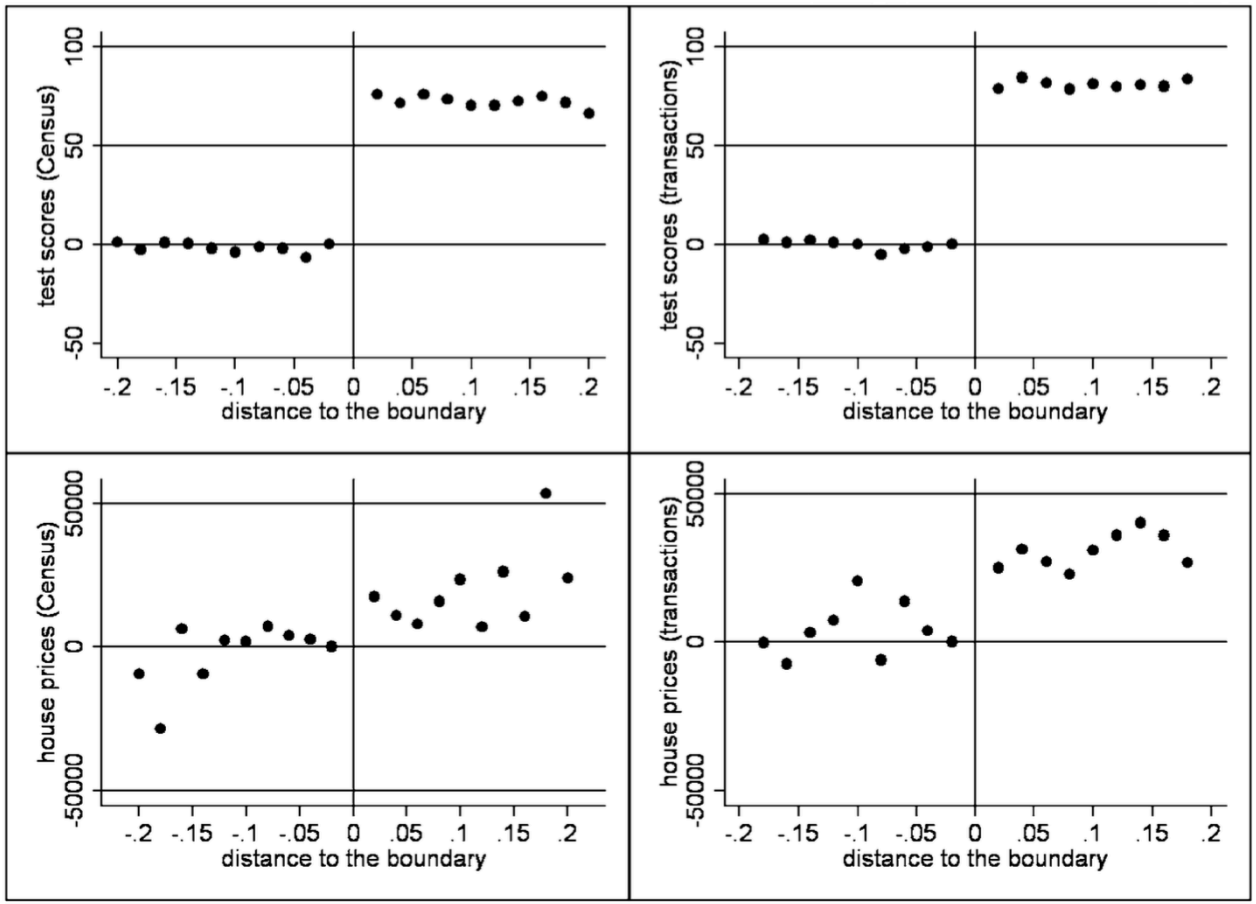


FIG. 1.—Test scores and house prices around the boundary. Each panel is constructed using the following procedure: (i) regress the variable in question on boundary fixed effects and on 0.02-mile band distance to the boundary dummy variables; (ii) plot the coefficients on these distance dummies. Thus a given point in each panel represents this conditional average at a given distance to the boundary, where negative distances indicate the low test score side.

Not so much for "fixed" house characteristics

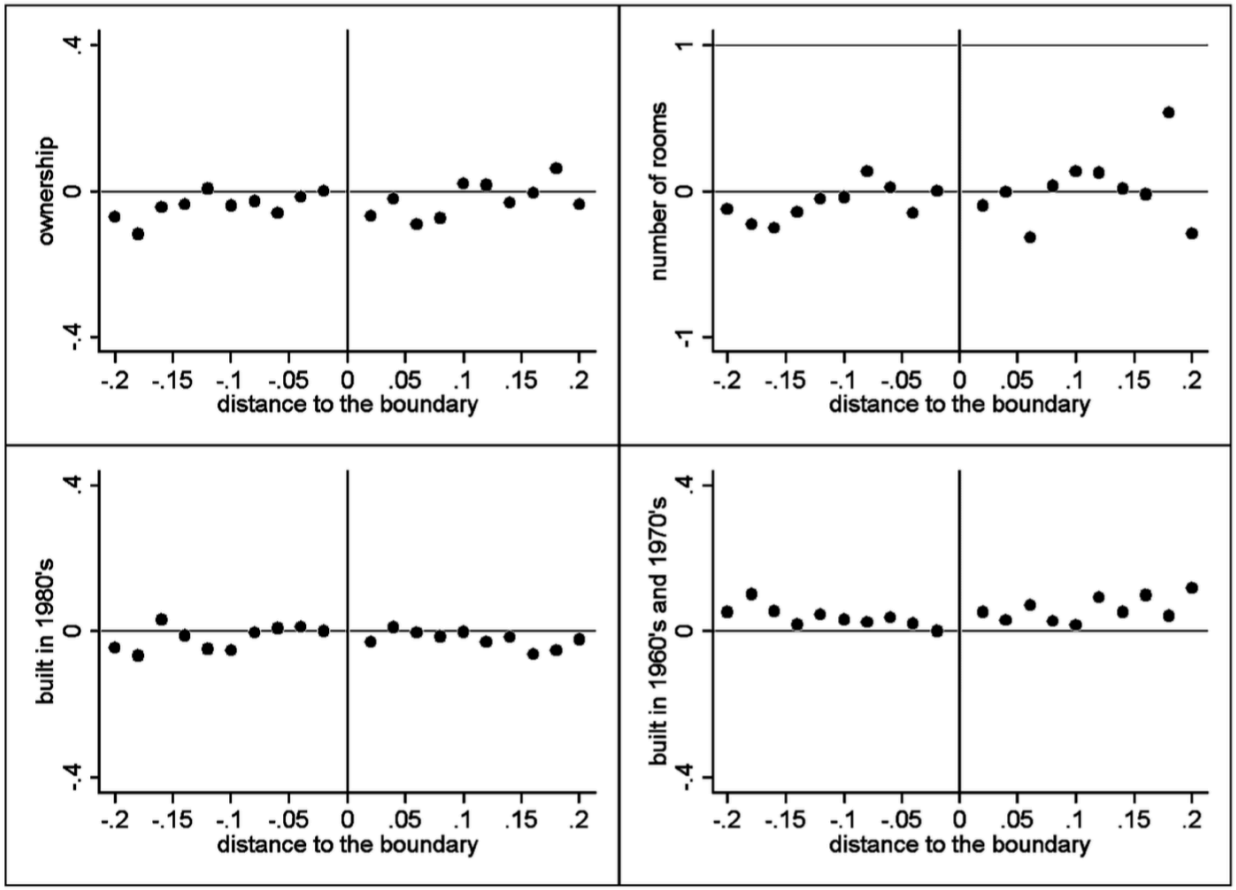


FIG. 2.—Census housing characteristics around the boundary. Each panel is constructed using the following procedure: (i) regress the variable in question on boundary fixed effects and on 0.02-mile band distance to the boundary dummy variables; (ii) plot the coefficients on these distance dummies. Thus a given point in each panel represents this conditional average at a given distance to the boundary, where negative distances indicate the low test score side.

The characteristics of neighbors is also discontinuous

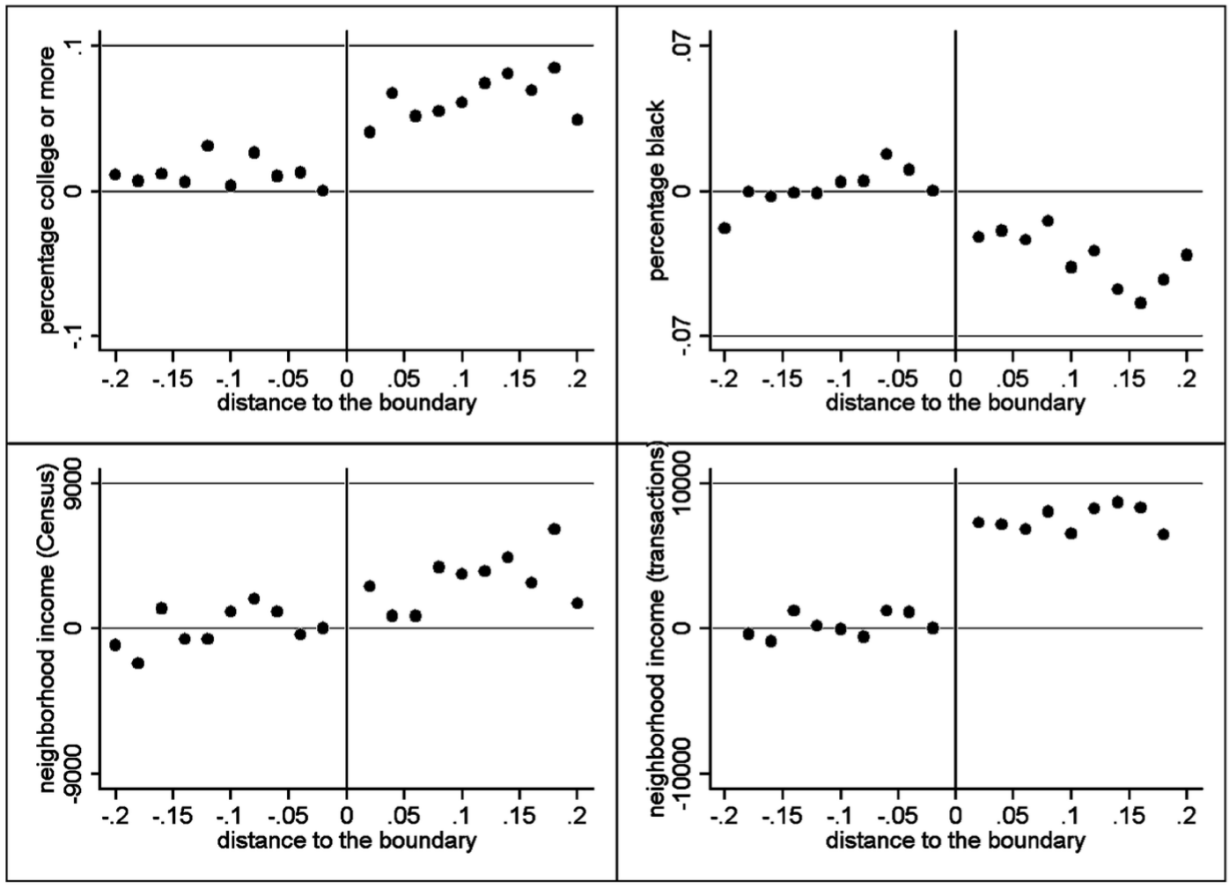


FIG. 4.—Neighborhood sociodemographics around the boundary. Each panel is constructed using the following procedure: (i) regress the variable in question on boundary fixed effects and on 0.02-mile band distance to the boundary dummy variables; (ii) plot the coefficients on these distance dummies. Thus a given point in each panel represents this conditional average at a given distance to the boundary, where negative distances indicate the low test score side.

Sorting based on sociodemographic characteristics?

Controlling the demographics of neighbors affects the hedonic estimate

TABLE 3
KEY COEFFICIENTS FROM BASELINE HEDONIC PRICE REGRESSIONS

	Sample			
	Within 0.20 Mile of Boundary (N = 27,548)		Within 0.10 Mile of Boundary (N = 15,122)	
Boundary fixed effects included	No	Yes	No	Yes
	A. Excluding Neighborhood Sociodemographic Characteristics			
	(1)	(2)	(5)	(6)
Average test score (in standard deviations)	123.7	33.1	126.5	26.1
	(13.2)	(7.6)	(12.4)	(6.6)
R^2	.54	.62	.54	.62
	B. Including Neighborhood Sociodemographic Characteristics			
	(3)	(4)	(7)	(8)
Average test score (in standard deviations)	34.8	17.3	44.1	14.6
	(8.1)	(5.9)	(8.5)	(6.3)
% census block group black	−99.8	1.5	−123.1	4.3
	(33.4)	(38.9)	(32.5)	(39.1)
% block group with college degree or more	220.1	89.9	204.4	80.8
	(39.9)	(32.3)	(40.8)	(39.7)
Average block group income (/10,000)	60.0	45.0	55.6	42.9
	(4.0)	(4.6)	(4.3)	(6.1)
R^2	.59	.64	.59	.63

NOTE.—All regressions shown in the table also include controls for whether the house is owner-occupied, the number of rooms, year built (1980s, 1960–79, pre-1960), elevation, population density, crime, and land use (% industrial, % residential, % commercial, % open space, % other) in 1-, 2-, and 3-mile rings around each location. The dependent variable is the monthly user cost of housing, which equals monthly rent for renter-occupied units and a monthly user cost for owner-occupied housing, calculated as described in the text. Standard errors corrected for clustering at the school level are reported in parentheses.

Neighborhood Sorting

- Identification/estimation strategy:
  - **Step 1:** Estimate the *average* willingness-to-pay of consumers for neighborhood quality, by accounting for heterogeneous tastes.
  - **Step 2:** Decompose the average willingness-to-pay into taste for school quality and neighborhood characteristics (i.e. peers).
- The exercise is analogous to Berry et al. (1995):
  - **Revealed preference:** The average WTP of a product is the quality index that rationalize aggregate choices (i.e. inverse demand).
  - **Unobserved attributes:** School quality is correlated with unobserved characteristics of a neighborhood (i.e. just like price)

Neighborhood Sorting Model

- Households choose which house to buy:

$$\max_{h \in H} V_h^i = X_h \alpha_x^i - \alpha_p^i p_h + \xi_h + \varepsilon_h^i$$

  - **Herogeneity:** $\alpha_j^i = \bar{\alpha}_j + \sum_{k=1}^{K} \sigma_{j,k} z_k^i$, where j indexes a house characteristic and k indexes a consumer attribute.
  - **Characteristics:** X_h includes average test-score, neighborhood demographic characteristics and other physical neighborhood attributes
- **Assumption:** Heterogeneity in taste is capture entirely by observed demographic characteristics (z^i) and a T1EV shock ε_h^i
- The utility can be decompose into common and idiosyncratic terms:

$$V_h^i = \delta_h + \mu_h(z_i; \Sigma) + \varepsilon_h^i$$

  where $\delta_h = X_h \bar{\alpha}_x - \bar{\alpha}_p p_h + \xi_h$ denotes the average quality of house h, and $\Sigma = \{\sigma_{j,k}\}$.
- **Data:** Cross-section of households living in the Bay-area from the U.S. census: demographics, house address, workplace location, house value, income, etc.

Estimation Strategy: Step 1

- **Maximum likelihood:** Multinomial logit with choice-specify intercepts. Let h_i denote the choice that household i made.

$$\max_{\Sigma,\delta} \sum_i \sum_h I(h_i = h) \log \sigma_h(X; \Sigma, \delta, z_i)$$

$$\leftrightarrow \quad \max_{\Sigma} \sum_i \sum_h I(h_i = h) \log \sigma_h(X; \alpha, \delta(\Sigma), z_i)$$

$$\text{s.t.} \quad \delta_h(\Sigma): \quad 1 = \sum_i \sigma_h(X; \alpha, \delta(\Sigma), z_i), \quad \forall h \text{ (FOC)}$$

- The constraint corresponds to a market-clearing condition:
  - $\delta_h(\Sigma)$ is the quality of each house such that demand equal supply:

$$\sum_i P_h^i = D_h(\delta) = S_h = 1$$

  - The existence and uniqueness of this inverse function follows from Berry et al. (1995)

$$\delta_h(\Sigma) = \sigma_h^{-1}(s, X; \Sigma), \text{ where } s = 1/N$$

Estimation Strategy: Step 2

- Conditional of the ML estimate, $\hat{\Sigma}$, the average quality of house h is:

$$\delta_h(\hat{\Sigma})(=\sigma_h^{-1}(X;\hat{\Sigma})) = X_h\bar{\alpha}_x - \bar{\alpha}_p p_h + \xi_h$$

- Simultaneity problem(s): ξ_h is correlated with
  - School quality
  - Neighbors sociodemographics
  - House price
- Identifying assumptions:
  - School boundary conditional independence:

$$E(\xi_h|X_h, b_h) = 0$$

  - Valid price instrument: w_h

Estimation Strategy: Step 2

- **Quality decomposition:** Boundary fixed-effect IV regression

$$\delta_h = X_h \bar{\alpha}_x - \bar{\alpha}_p p_h + \mu_{b(h)} + \xi_h$$

  where $\mu_{b(h)}$ is a school boundary fixed-effect.
- **Price IV:** (w_h)
  - Physical characteristics of neighboring houses (initial estimate/weak IV)
  - Market clearing prices if all endogenous attributes would be zero (i.e. $\xi_j = 0$ and characteristics of neighbors).
- **Endogenous characteristics:** Fixed-effect approach
  - This is less convincing...
  - Assumption: $E(\xi_h | \mu_{b(h)}, X_h, w_h) = 0$.
  - This would be consistent with a model in which the "stock" of neighbors is fixed in the short-run, and determined before ξ_h is realized.
  - Which is not really consistent with the model: The market clearing condition holds for every household.

Link with Hedonic Regression

- The quality decomposition regression can be written in terms of inverse-demand:

$$\delta_h(\hat{\Sigma}) = X_h\bar{\alpha}_x - \bar{\alpha}_p p_h + \xi_h \leftrightarrow p_h = X_h\tilde{\alpha}_x - \frac{1}{\bar{\alpha}_p}\sigma_h^{-1}(X;\hat{\Sigma}) + \tilde{\xi}_h$$

- The third term, $\sigma_h^{-1}(X;\hat{\Sigma})$, is an adjustment due to sorting based on heterogenous preferences
  - **If $\Sigma = 0$:** σ_h^{-1} is a constant (i.e. logit), and we're back to the standard hedonic regression

$$p_h = X_h\tilde{\alpha}_x + \tilde{\xi}_h, \quad \tilde{\alpha}_x = \bar{\alpha}_x/\bar{\alpha}_p$$

  - **If $\Sigma \neq 0$:** The hedonic regression identifies the WTP of the *marginal* consumer (not the average), and the boundary fixed-effect estimate of $\bar{\alpha}$ is biased.

$$p_h = X_h\tilde{\alpha}_x + \underbrace{\frac{-1}{\bar{\alpha}_p}\sigma_h^{-1}(X;\Sigma) + \tilde{\xi}_h}_{\text{Residual}}$$

For the most part the heterogeneous sorting results are consistent with the hedonic (except for race)

TABLE 7
DELTA REGRESSIONS: IMPLIED MEAN WILLINGNESS TO PAY
SAMPLE: WITHIN 0.20 MILE OF BOUNDARY (N = 27,458)

Boundary fixed effects included	No	Yes
	A. Excluding Neighborhood Sociodemographic Characteristics	
	(1)	(2)
Average test score (in standard deviations)	97.3	40.8
	(14.0)	(5.5)
	B. Including Neighborhood Sociodemographic Characteristics	
	(3)	(4)
Average test score (in standard deviations)	18.0	19.7
	(8.3)	(7.4)
% block group black	−404.8	−104.8
	(41.4)	(36.9)
% census block group Hispanic	−88.4	−3.5
% block group with college degree or more	183.5	104.6
	(26.4)	(31.8)
Average block group income (/10,000)	30.7	36.3
	(3.7)	(6.6)

NOTE.—All regressions shown in the table also include controls for whether the house is owner-occupied, the number of rooms, year built (1980s, 1960–79, pre-1960), elevation, population density, crime, and land use (% industrial, % residential, % commercial, % open space, % other) in 1-, 2-, and 3-mile rings around each location. The dependent variable is the monthly user cost of housing, which equals monthly rent for renter-occupied units and a monthly user cost for owner-occupied housing, calculated as described in the text. Standard errors corrected for clustering at the school level are reported in parentheses.

Heterogeneity in preferences: People prefer to self-segregate based on race and education

TABLE 8
HETEROGENEITY IN MARGINAL WILLINGNESS TO PAY FOR AVERAGE TEST SCORE AND NEIGHBORHOOD SOCIODEMOGRAPHIC CHARACTERISTICS

		Neighborhood Sociodemographics		
	Average Test Score +1 SD	+10% Black vs. White	+10% College-Educated	Block Group Average Income +$10,000
Mean MWTP	19.69	−10.50	10.46	36.3
	(7.41)	(3.69)	(3.18)	(6.60)
Household income (+$10,000)	1.38	−1.23	1.41	.86
	(.33)	(.37)	(.21)	(.12)
Children under 18 vs. no children	7.41	11.86	−16.07	2.37
	(3.58)	(3.03)	(2.25)	(1.17)
Black vs. white	−14.31	98.34	18.45	−1.16
	(7.36)	(3.93)	(4.52)	(2.24)
College degree or more vs. some college or less	13.03	9.19	58.05	.31
	(3.57)	(3.14)	(2.33)	(1.40)

NOTE.—The first row of the table reports the mean marginal willingness to pay for the change reported in the column heading. The remaining rows report the difference in willingness to pay associated with the change listed in the row heading, holding all other factors equal. The full heterogeneous choice model includes 135 interactions between nine household characteristics and 15 housing and neighborhood characteristics. The included household characteristics are household income, the presence of children under 18, and the race/ethnicity (Asian, black, Hispanic, white), educational attainment (some college, college degree or more), work status, and age of the household head. The housing and neighborhood characteristics are the monthly user cost of housing, distance to work, average test score, whether the house is owner-occupied, number of rooms, year built (1980s, 1960–79, pre-1960), elevation, population density, crime, and the racial composition (% Asian, % black, % Hispanic, % white) and average education (% college degree) and household income for the corresponding census block group. Standard errors are reported in parentheses.

Hedonic approach to discrete-choice models

Reference: Bajari and Benkard (2005)

- Extends the hedonic methodology:
  - Relax the perfect competition assumption (original Rosen model)
  - Adopt a 2-step method to estimate the distribution of marginal utilities for discrete and continuous characteristics.
  - Applicable in settings with incomplete product space (i.e. bundles)
- Methodological framework that removes the idiosyncratic error (i.e. logit)
- Pure characteristic model of demand:

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

See also Berry and Pakes (2007)

Model description

- Consumer's Problem:

$$\max_{j,c} u_i(x_j, \xi_j, c), \quad s.t. \quad p_{jt} + c = y_i$$

- Quasi-linear example: $u_i = \log(x_j)\beta_{i,x} + \log(\xi_j)\beta_{i,\xi} + (y_i - p_{jt})$
- **Hedonic price function:** $p(x_j, \xi_j)$ have the following properties:
  1. There exists one price for each bundle of characteristics
  2. p_j is increasing in ξ_j
  3. p_j is Lipschitz-continuous (i.e. continuous and bounded).
- **Question:** What assumptions on $u_i(\cdot)$ do we need to get properties (1)-(3)?

Model description

- **Assumptions:**
  1. $u_i()$ is continuously differentiable in c and strictly increasing in c
  2. $u_i()$ is is Lipschitz-continuous in (x_j, ξ_j)
  3. $u_i()$ is strictly increasing in ξ_j (no-satiation).
- Implications:
  - If two goods have the same x_j and the same ξ_j, they must have the same price (otherwise one has zero demand)
  - If $x_j = x_{j'}$ and $\xi_j > \xi_{j'}$ then $p_j > p_{j'}$, and (iii) $|p_j - p_{j'}| \leq M(|x_j - x_{j'}| + |\xi_j - \xi_{j'}|)$, for $M < \infty$
- **Key result:** There exists a continuous and differentiable pricing function $p(x_j, \xi_j)$.

Identification: First-stage

- **Goal:** Identify ξ_j and $p_t(x_j, \xi_j)$ from a cross-section of products up to a normalization: $\xi_j \in [0, 1]$.
- **Identification 1:** ξ is independent of x (i.e. like in BLP)

$$\begin{aligned} F_{p_t|x=x_j}(p_{jt}) &= \Pr(p_t(x, \xi) < p_{jt}|x = x_j) \quad \text{Data} \\ &= \Pr(\xi \leq p_j^{-1}(x, p_{jt}|x = x_j) \quad \text{Continuous} \\ &= \Pr(\xi \leq p_j^{-1}(x_j, p_{jt})) \quad \text{Independence} \\ &= \xi_j \quad \text{Normalization} \end{aligned}$$

- In practice: Estimate $F_t(p_{jt}, x_{jt})$ using flexible regression methods (e.g. local polynomials).

Relaxing exogenous product characteristics

- **Option packages:** Estimate a brand fixed-effect $\xi_j = F_{x=\bar{x}}(p_{jt}, x_{jt})$ for all product categories using a "baseline" package that is common across brands.
- **IV approach:** There exists $x_1 = h(z, x_{-1}, \eta)$ where $E(\xi|x, \eta) = E(\xi|\eta)$ (i.e. x's are conditionally independent of ξ).
- Non-parametric IV regression:
  1. Recover $\eta = F_t(x_1|x_{-1}, z)$
  2. Recover ξ_j using the distribution of η:

$$\xi_j = \int \Pr(\xi < p^{-1}(x_j, p_{jt})|\eta) dF(\eta)$$

Second-stage: Estimation of preferences

- **Continuous Xs:** If consumers are price-takers, the optimal bundle of characteristics generates a vector of FOCs:

$$\partial u_i/\partial x_j - \partial p_j/\partial x_j = 0$$

$$\beta_{i,x}\frac{1}{x_{ij}} - \partial p_j/\partial x_{ij} = 0$$

- Therefore, under quasi-linear utility, we can directly estimate the distribution of preference parameters from consumer-level data on product choices.
- **Discrete choice-set:** Optimality conditions is a set of inequalities:

$$u(x_j, p_j, \xi_j|\beta_i) \geq u_k(x_k, p_k, \xi_k|\beta_i) \quad \forall \neq j.$$

- If we assume a parametric distribution for β_i, we can use the inequalities to form a likelihood function, or moment inequality conditions.

Some application results from the PC market

TABLE 5
DISTRIBUTION OF BOUNDS ON WELFARE LOSS FROM FACING AUGUST 1997 CHOICE SET
(%)

	Percentile							Average (Excluding ∞'s)
	.05	.10	.25	.5	.75	.9	.95	
	Model 1							
Lower bound	7	9	10	11	12	13	14	11
Upper bound	10	11	12	13	15	19	26	16
	Model 2							
Lower bound	−7	−2	4	11	15	17	19	9
Upper bound	21	23	25	31	43	67	94	42

TABLE 6
DISTRIBUTION OF BOUNDS ON WELFARE LOSS FROM REMOVING EACH PRODUCT ($)

	Percentile							Average (Excluding ∞'s)
	.05	.10	.25	.5	.75	.9	.95	
	Model 1							
Lower bound	0	0	0	0	0	0	0	0
Upper bound	0	0	1	3	9	26	64	15
	Model 2							
Lower bound	0	0	0	0	0	0	0	0
Upper bound	6	13	27	59	177	526	851	213

Multiple discrete choice and demand for computers

Reference: Hendel (1999)

- Measure the welfare gains from computerization in banking and insurance
- How?
  - Traditional papers have used a production function approach: Measure the gains in productivity from computers. Mixed results.
  - This paper uses a revealed-preference approach, that takes into account: (i) quality differentiation, and (ii) intensive margin (multiple-brand/multipe-units).
  - Estimate demand for computers at the firm level: (i) durable good, (ii) buyers purchase more than one computer, (iii) firms buy more than one variety.
  - Proposed explanation:
    - ⋆ Assume away possible complementarities ($\neq$ Gentzkow)
    - ⋆ Each firm is composed of J_f **tasks**, each of which have different computational needs.
    - ⋆ Firms differ in the number of tasks.

Data and Notation

- Data: (i) price and characteristics for each model in 1988, (ii) firm level data on the stock of computers of each kind.
- Key variables:
  - Each computer has N characteristics + Brand fixed-effect (I): C_i.
  - Firms are differentiated in terms of their "demographic" attributes: D_f
  - The number of tasks differ across firms:

$$J_f \sim \Gamma(g' D_f) \text{ (poisson distribution)}$$

  - Choice variables: X_j = number and brand of computers for task j.

Profit Function

- Profit of the firm is separable in each task:

$$\Pi^f(X_1, ..., X_{J_f}) = \sum_{j=1}^{J_f} \pi_j^f(D_f, A_j, X_j)$$

  where A_j is a vector of taste shock for task j.

- Functional form:

$$\pi_j^f = \left(\sum_i W_{ijf} x_{ijf} \right)^{\alpha} S(D_f) - \sum_i p_i x_{ijf}$$

- Where,
  - $S(D_f)$ is computer productivity of firm f
  - $W_{ijf} = (\max(0, A_j C_i))^{m(D_f)}$ is the valuation of computer i for task j,
  - $\max(0, A_j C_i)$ is the quality of computer i from task's j point of view
  - $m(D_f)$ is the **return** on quality for firm j

Profit Function: Interpretation

- **Differentiation:**
  - $m(D_f)$ measures the degree of vertical differentiation in the market. If $m(D_f) = m$ for all firms, every firm values quality the same way, and the market is not vertically differentiated.
  - The random coefficients A_j creates horizontal differentiation: Each task has a different ranking of computer quality.
  - Taste shocks A_j are normally distributed.
- **Durable good?**
  - Prices are adjusted for depreciation: $p_i = P_{it} - E(P_{i,t+1}) = (1-\delta)P_{it}$ and $\delta = 0.25$. For old products, assume larger depreciation (i.e. d_0).
- **Implications:**
  - $0 < \alpha < 1$: firms consume a finite number of computers,
  - Since the quality is additive, each task will consumer at most one variety (not-testable),

Model Solution

- For a given vector of taste shock and tasks:
  1. Solve the continuous approximation to the task problem (analytical)
  2. Compute the reduced form profits (analytical)
  3. Round the number of computers and use the highest profit
  4. Store $i_j^* = \max_i \pi_i^*$ and X_{ijf}^*
- This process is repeated for a large number of simulation draws of the taste shocks

Estimation method: GMM

- Aggregation of computer demands at the firm level:

$$X_f(D_f, \theta) = \sum_0^{\infty} \left[\sum_0^{J_f} \int X_{jf}^*(D_f, A, \theta) \mu(dA|D_f, \theta) \right] P(dJ|D_f, \theta)$$

- Integrals are approximated via simulation: Independent draws for tasks and taste shocks.
- Moment conditions:

$$\frac{1}{F} \sum_f T(D_f) \otimes \left(X_f(D_f, \theta_0) - \hat{X}_f \right) \approx 0$$

Summary of the Results

- Benefits from computerization: $\Pi_f(D_f) \approx 1,16$ Billlions.
- Profit by computer is on average $1459 < 1587$ for the average price
- Simulated markups are quite low. Except for IBM which sells higher quality products.

Parameter estimates

	Coefficient	Std. Dev.
s_0	137·5	16·63
$s_{0\text{ins}}$	16·32[a]	36·16
s_1	0·46	0·12
m_1	0·0004	0·00011
m_{ins}	0·005[a]	4·90
α_0	0·28	0·01
Var (A_i)	11·47	0·88
Var (A_m)	2·43	0·38
g_1	1·65	0·60
g_2	1·93	0·18
d_0	0·76	0·05
$G_F(\Theta) = 38·76$		

Key:
s_0–intercept of the function $S(D_f)$.
s_1–slope of $S(D_f)$ with respect to emp_f.
s_{ins}–coefficient of the insurance dummy on $S(D_f)$.
m_1–slope of $m(D_f)$ with respect to wh_f.
m_{ins}–coefficient of the insurance dummy on $m(D_f)$.
α_0–profit function exponent.
A_i–mean valuation of make/model i.
A_m–mean valuation for MHz.
g_1–slope of task generating process mean w.r.t. emp_f.
g_2–slope of task generating process mean w.r.t. soft_f.
d_0–coefficient on dp_f.
/a–non-significant coefficient.
To compute the variance matrix I used numerical

TABLE IV

Demand price elasticities

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. IBM 8086	–9·0	0·4	0·4	0	0·5	3·2	1·4	0	1·2	0·3	1·7	2·2	1·8	4·8
2. IBM 8088	1·6	–9	1·7	0·4	1·9	0·7	2·2	35	1·5	1·1	2·1	2·5	1·5	2·2
3. IBM 286	3·6	5·3	–9·0	4·1	9·3	5·9	11	50	9·0	13	4·2	6·9	10	4·6
4. IBM 386	0·4	1·0	2·8	–8·0	6·2	1·2	0·4	26	12	29	0·1	0·4	5·8	0·1
5. Apple 68000	0·6	0·7	0·9	0·8	–17	8·9	1·4	0	0·8	1·0	0·7	1·0	0·6	0·4
6. Apple 68020	0·1	0	0	0	0·2	–20	0	0	0	0	0	0	0·1	0
7. Zenith 8088	0·1	0·1	0·1	0	0·1	0	–18	0	0·5	0	0·1	0·2	0·3	0·5
8. Zenith 386	0	0	0	0	0	0	0	–23	0	0	0	0	0	0
9. Compaq 286	0·2	0·1	0·1	0	0·1	0	0·7	0	–15	0·1	0·1	0·3	0·5	0·4
10. Compaq386	0·1	0·1	0·2	0·7	0·1	0	0	68	0·2	–38	0	0·1	0·8	0·1
11. Clone 8088	0·3	0·1	0·1	0	0·1	0	0·5	0	0·3	0·2	–10	0·5	0·5	0·6
12. Clone 286	0·8	0·2	0·3	0	0·3	0·1	1·4	0	0·9	0·4	1·3	–14	1·7	2·0
13. Clone 386	0·7	0·2	0·3	0·4	0·4	2·0	1·4	45	1·0	1·3	0·7	1·4	–26	1·1
14. AT&T 8086	0·1	0	0	0	0	0	0·3	0	0	0	0·2	0·2	0·2	–14
Markups (%)	14·8	16·2	19·2	19·7	6·3	5·4	5·8	4·6	7·4	2·7	11·6	8·6	4·1	7·8

Columns represent demand elasticities of a specific model with respect to the prices of the models in the first column. Elasticities are computed using a numerical derivative w.r.t. price. Part of the elasticities reported as zero are nonzero although negligible. Notice that simulated demand is non-differentiable for any finite number of simulation draws. That explains part of the reported zeros.

Valuing new goods in a model with complementarity

Reference: **?**)

- **Motivation:**
  - Traditional discrete choice model assumes strict substitutability between goods, even when they are consumed together.
  - Identification problem: selection versus true complementarity.
  - Application: Online versus paper news.
- **Example:** 2 good + outside option
  - Four possible packages:

$$
\begin{aligned}
u_0 &= 0 \\
u_A &= \delta_A - \alpha p_A + \nu_A \\
u_B &= \delta_B - \alpha p_B + \nu_B \\
u_{AB} &= u_A + u_B + \Gamma
\end{aligned}
$$

  - Where $(\nu_A, \nu_B)^T \sim N\left(0, \begin{pmatrix} 1 & \sigma \\ \sigma & 1 \end{pmatrix}\right)$.

Complements vs Subsitutes

- In this context, demand for good A is:

$$Q_A = P_A + P_{AB}$$

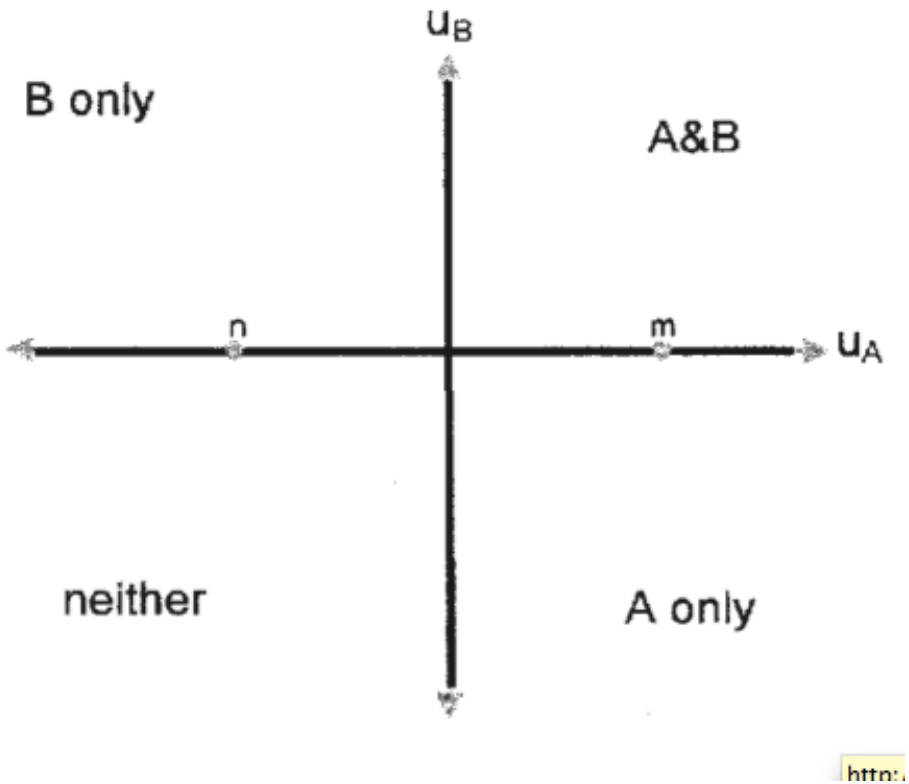
where:

$$P_{AB} = \int 1(u_{AB} > 0)1(u_{AB} > u_A, u_{AB} > U_B)dF(\mathbf{u})$$

$$\text{and } P_j = \int 1(u_j < 0)1(u_j > u_{AB}, u_j > u_{-j})dF(\mathbf{u}), \forall j = \{A, B\}$$

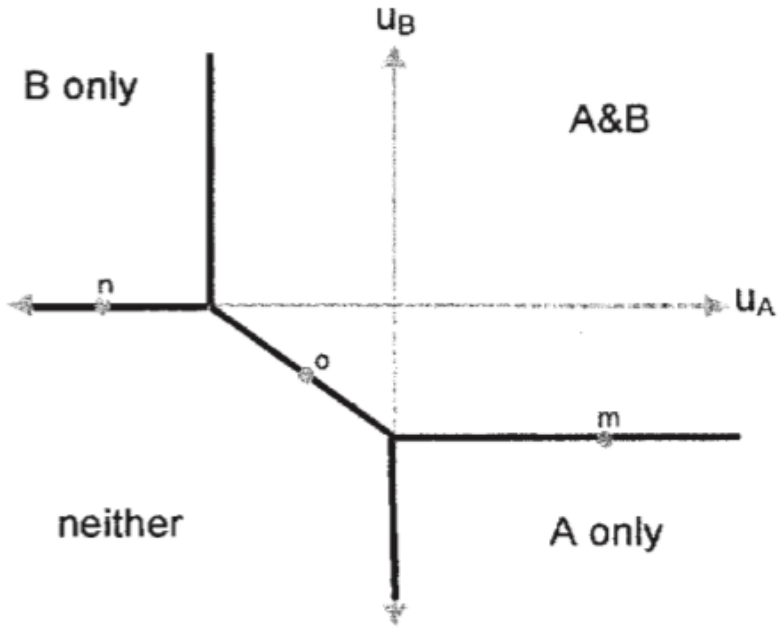
- **Definition:** A is complement with B if $\partial Q_A/\partial p_B < 0$.

Independent case: $\Gamma = 0$



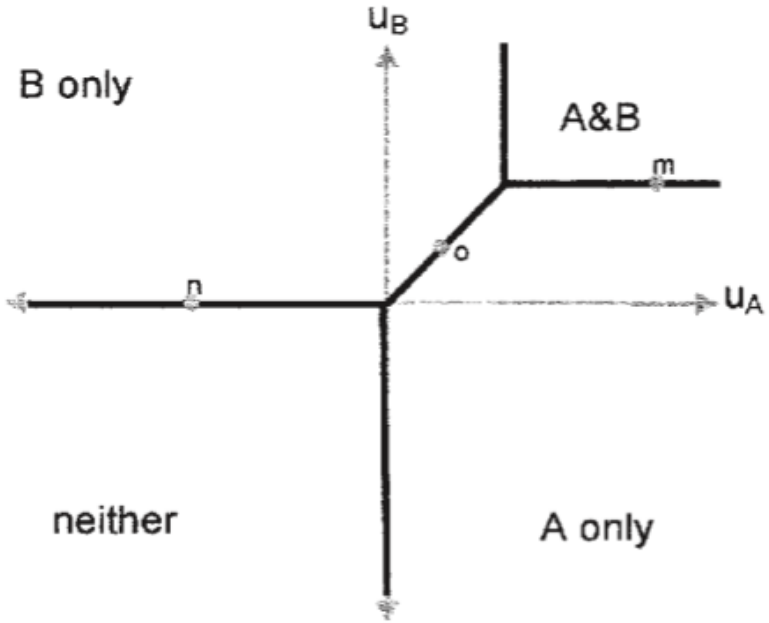
- Thought experiment: After an increase in p_B:
  - n-type: switch out of B to 0
  - m-type: switch out of AB to A
  - Therefore, $|\Delta \Pr_{AB}| = |\Delta \Pr_A| = |\Delta \Pr_B|$, and $\partial Q_A / \partial p_b = 0$

Complement case: $\Gamma > 0$



- Thought experiment: After an increase in p_B:
  - n-type: switch out of B to 0
  - m-type: switch out of AB to A
  - o-type: switch out of AB to 0
  - Therefore, $|\Delta \Pr_{AB}| > |\Delta \Pr_{A}|$, and $\partial Q_A / \partial p_b < 0$

Substitute case: $\Gamma > 0$



- Thought experiment: After an increase in p_B:
  - n-type: switch out of B to 0
  - m-type: switch out of AB to A
  - o-type: switch out of B to A
  - Therefore, $|\Delta \Pr_{AB}| < |\Delta \Pr_{A}|$, and and $\partial Q_A / \partial p_b > 0$

Identification of 2-good model: $(\delta_A, \delta_B, \alpha, \Gamma, \sigma)$

- Baseline moments: Pr_{AB}, Pr_A, Pr_B.
- α: Can be identified from exogenous price variation, and/or supply-side restriction (here)
- Without further assumptions, Γ and σ cannot be separately identified. Why?
  - Moment accounting...
  - Assume: $\text{Pr}_{AB} > 0$.
  - Standard selection problem: (i) unobserved heterogeneity, or (ii) complementarity.

Solutions

- Exclusion restriction:
  - X affects only the value of A directly (e.g. internet at work for online newspaper).
  - Creates over-identification moments restrictions: $\Pr_{AB}(X)$, $\Pr_{A}(X)$, $\Pr_{B}(X)$.
  - Observing an increase in u_B we can measure the relative change in $\Pr_A$ and $\Pr_{AB}$:
    - $\Gamma = 0$: $\Delta \Pr_{AB} = \Delta \Pr_A$
    - $\Gamma > 0$: $\Delta \Pr_{AB} > \Delta \Pr_A$
    - $\Gamma < 0$: $\Delta \Pr_{AB} < \Delta \Pr_A$
- Panel data: Usual argument (i.e. assuming persistence in ν_i).

Application: Demand for Online Newspaper

- Micro-data: Survey of 16,176 adults between 2000-2003.
- Key variables:
  - All newspaper read over the last day, and last 5 years ($\sim$ panel)
  - Demographics + Internet access
  - No price variation: (i) online is free, and (ii) subscription is constant.
  - Products: Post, Times, Post.com, Other. (i.e. bundle outside option for news + no news at all)
- Reduced-form evidence:
  - 48% read some news
  - 28% read more than one paper
  - The fraction of online readers among "offline" readers is larger than among non-news reader (i.e. news-junkies).
  - The positive correlation holds ever after controlling for X's.
- IV-strategy:
  - Online shifter: internet at work, fast internet access, internet for education or work.

Reduced-form Results

	OLS	IV		
		(1)	(2)	(3)
Dependent variable: Read Post *last 5 days*				
Read post.com last 5 days	0.0464**	−0.4132**	−0.4579**	−0.4381**
	(0.0090)	(0.107)	(0.119)	(0.141)
Other Internet news				0.0244
				(0.0195)
Industry controls			X	
Occupation controls			X	
Overidentification test p value		0.302	0.431	0.219
R-squared	0.333	0.214	0.207	0.202
N	14313	14313	14313	14313

Notes: Robust standard errors are in parentheses. The first row gives coefficients on a dummy for reading the post.com in the last five weekdays. IV regressions instrument for post.com consumption with dummy variables for Internet access at work, fast Internet connection, and reported use of the Internet for research/education and work-related tasks. Overidentification test p value is the p value from a standard Sargan test. Other Internet news is a dummy for online news use other than online newspapers. Industry controls are dummies for 12 industry categories. Occupation controls are dummies for 11 occupation categories. All regressions include controls for *Washington Times* readership, age, sex, education (four categories), white-collar work, computer work, employment status, income, political party, date of survey, location of residence within the DMA (six categories), and dummy variables for the number of missing values. The regressions omit observations where print newspapers are not in the choice set (consumer reports that she generally reads no newspaper sections) and control for presence of online newspapers in the choice set (whether the consumer used the Internet in the last 30 days).

** Significant at 1 percent.

Structural Model

- Product and bundle utility:

$$u_{ijt} = \delta_j + x_i \beta_j + \nu_{ij} + \tau_{it}$$

$$U_{irt} = \sum_{j \in r} u_{ijt} + \Gamma_r + \epsilon_{irt}$$

- Price effect: $\delta_j = -\alpha p_j + \xi_j$ where $r \neq 0$ is the bundle identifier, $U_{irt} = \epsilon_{irt}$, and β_j involve exclusion restrictions.
- Assumptions:
  - $\epsilon_{irt} \sim$ T1EV$(0, 1)$, and $\nu_i \sim N(0, \Sigma)$
  - $\tau_{it} \in \{0, \tau\}$ with probability $(1 - \gamma)$ and γ. (to match the panel)

Likelihood Function

- Choice probability:
  - 24 hours period:

$$Q_r(X_i, \nu_i, \tau_i|\theta) = \frac{\exp(U_{ir})}{1 + \sum_{r'} \exp(U_{i,r'})}$$

    where r is a bundle.
  - 5-days period: Let q be the observed choices over 5 days, and $\Omega(q)$ be the set of possible sequences of bundle choices.

$$P_q(X_i, \nu_i|\theta) = \sum_{\tilde{q} \in \Omega(q)} \prod_{t=1}^{5} [\gamma Q_{i,\tilde{q}_t}(X_i, \nu_i, \tau|\theta) + (1-\gamma) Q_{i,\tilde{q}_t}(X_i, \nu_i, 0|\theta)]$$

- Likelihood contribution:

$$L_i(q_i, X_i|\theta) = \int P_{q_i}(X_i, v_i|\theta) f(v_i; \Sigma) dv_i \approx \frac{1}{S} \sum_s P_{q_i}(X_i, v_s|\theta)$$

where $\{v_1, \ldots, v_S\}$ is a sequence of random vectors from $f(v; \Sigma)$.

Estimation Results

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)		

Notes: Standard errors in parentheses. The model also includes a third-order interaction term for the *Post-Times*-post.com bundle which is not reported in the table. Fast connection indicates consumers with DSL, cable modem, or T1 connections at home. Use variables were responses to the question, "In what ways do you use online services?"

** Significant at 1 percent.

Estimation Results

TABLE 7—VARIANCE AND COVARIANCE OF CONSUMER CHARACTERISTICS

Covariance of observable utility			
	Post	post.com	*Times*
Post	14.0		
post.com	7.95	8.43	
Times	7.66	4.85	10.2
Covariance of ν unobservables			
	Post	post.com	*Times*
Post	11.0		
post.com	4.14	23.5	
Times	−1.17	5.19	86.1
Variance of τ unobservables			2.03
Variance of ε unobservables			1.64

Note: The table shows the covariance of different components of utility at the estimated parameter values.

Price Effects and the Value of Online News

- **Challenge:** The price of newspaper does vary in the data, and is zero for online options.
- **Solution:** Use a supply-side restriction to infer α_p from δ_{pos}.
- Simplify profit function for the Washington Post:

$$\Pi = \left[\text{Ad. Revenue}_p + \frac{p - c_p}{\text{Nb. reader per copy}}\right] N_p + \text{Ad. Revenu}_w N_w - \text{FC}$$

- Advertising revenue per reader and print marginal cost are estimated using reported balance sheet data.
- The price coefficient is estimated from the FOC:

$$-\alpha \left[\text{Ad. Revenue}_p + \frac{p - c_p}{\text{Nb. reader per copy}}\right] \frac{\partial N_p}{\partial \delta_{post}} + \text{Ad. Revenu}_w \frac{\partial N_w}{\partial \delta_{post}} = 0$$

Estimation Results

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)

Notes: Standard errors in parentheses. The table shows three measures of the online edition's impact. The cross-price derivative is the change in post.com readership when the *Post*'s price is increased by $.10. Change in print readership and print profits are the total changes for the *Post* when the online edition is added to the choice set. The table shows the estimated values in three models. Case 1 is the estimates from the full model. Case 2 is a model with observable consumer characteristics but no unobservables other than the i.i.d. logit errors. Case 3 is a model with no observable or unobservable consumer heterogeneity except the i.i.d. logit errors.

Bajari, P. and L. Benkard (2005).
Demand estimation with heterogeneous consumers and unobserved product characteristics: A hedonic approach.
Journal of Political Economy 113(6), 1239–1276.

Bayer, P., F. Ferreira, and R. McMillan (2007).
A unified framework for measuring preferences for schools and neighborhoods.
Journal of Political Economy 115(5), 588–638.

Berry, S., J. Levinsohn, and A. Pakes (1995).
Automobile prices in market equilibrium.
Econometrica: Journal of the Econometric Society 63(4), 841–890.

Berry, S. T. and A. Pakes (2007).
The pure characteristics demand model.
International Economic Review 48(4), 1193–1225.

Hendel, I. (1999, Apr.).
Estimating multiple-discrete choice models: An application to computerization returns.

Review of Economic Studies 66(2), 423–446.