Differentiated Products Demand Systems (B)

Jonathan Levin

Economics 257
Stanford University

Fall 2009

Demand in characteristic space: introduction

- Theory can be divided to two:
  - Price competition, taking products as given (see Caplin and Nalebuff, 1991, who provide conditions for existence for a wide set of models)
  - Competition in product space with or without subsequent price competition (e.g. Hotelling on a line, Salop on a circle, etc.).
- The empirical literature is almost entirely focused on the former, and there is much room for empirical analysis of the latter.
- Moreover, much of the demand literature uses the characteristics as instruments. This is both inefficient (why?) and probably inconsistent (why?); we all recognize it, but keep doing it without good alternatives (we will come back to it later).

Characteristic space: overview

- Products are bundles of characteristics, and consumers have preferences over these characteristics.
- Typically, we use a discrete choice approach: consumers choose one product only. Different consumers have different characteristics, so in the aggregate all products are chosen.
- Aggregate demand depends on the entire distribution of consumers.

Characteristic space: overview

- Formally, consumer i has the following utility from product j:

$$U_{ij} = U(X_j, p_j, \nu_i; \theta)$$

  We typically think of $j = 0, 1, 2, ..., J$, where product 0 is the outside good (why do we need it?).

- Consumer i's choice is the product which maximizes her utility, i.e. she chooses product j iff $U_{ij} \geq U_{ik}$ for all k. She chooses only one unit of one product, by assumption (how bad is this assumption?).
- Predicted market share for product j is therefore

$$s_j(\theta) = \int I\left(\nu_i \in \{\nu | U(X_j, p_j, \nu; \theta) \geq U(X_k, p_k, \nu; \theta) \forall k\}\right) dF(\nu_i)$$

- Note: utility is invariant to monotone transformations, so we need to normalize. Typically: set $U_{i0} = 0$ and fix one of the parameters or the variance of the error.

Characteristic choice: examples

- Two goods: $j = 0, 1, 2$. $U_{ij} = \delta_j + \epsilon_{ij}$ (and $U_{i0} = 0$).
- Hotelling with quadratic transportation costs:

$$U_{ij} = \overline{u} + (y_i - p_j) + \theta d^2(x_j, \nu_i)$$

- Vertical model: $U_{ij} = \delta_j - \upsilon_i p_j$ ($\upsilon_i > 0$). What makes it vertical? example: first class, business, economy.
- Logit:

$$U_{ij} = \overline{u} + (y_i - p_j) + \delta_j + \epsilon_{ij}$$

  where the ϵ's are distributed extreme value i.i.d across i and j ($F(x) = e^{-e^{-x}}$). It looks like normal, but with fatter tails.
  - A key feature of this distributional assumption is that it gives us a closed-form solution for the integral over the max.

Characteristic choice (cont.)

- In general, we can classify the models into two main classes:
  1. $U_{ij} = f(y_i, p_j) + \delta_j + \sum_k \beta_k x_{jk} \nu_{ik}$ (Berry and Pakes, 2002, "Pure Hedonic") or
     $U_{ij} = f(y_i, p_j) + \delta_j + \sum_k \alpha_k (x_{jk} - \nu_{ik})^2$ (Anderson, de Palma, and Thisse, 1992: "Ideal Type"), with $f_y > 0$, $f_p < 0$, $f_{py} \geq 0$.
  2. $U_{ij} = f(y_i, p_j) + \delta_j + \sum_k \beta_k x_{jk} \nu_{ik} + \epsilon_{ij}$ (Berry, Levinsohn, and Pakes, 1995)
- The key difference is the ϵ_{ij}. With the ϵ_{ij} the product space can never be exhausted: each new product comes with a whole new set of ϵ_{ij}'s, guaranteeing itself a positive market share and some market power. This may lead to problematic results in certain contexts, such as the analysis of new goods.
- Instruments: typically we assume X is exogenous, so we use instruments that are either cost shifters or functions of X which are likely to be correlated with markups.

The vertical model

- Utility is given by

$$U_{ij} = \delta_j - \upsilon_i p_j \ (\upsilon_i > 0)$$

  So if $p_j > p_k$ and $q_j > 0$, we must have $\delta_j > \delta_k$.
- Therefore, we order the products according to their price (and quality), say in an increasing order.
- Consumer i prefers product j over $j+1$ iff $\delta_j - \upsilon_i p_j > \delta_{j+1} - \upsilon_i p_{j+1}$ and over $j-1$ iff $\delta_j - \upsilon_i p_j > \delta_{j-1} - \upsilon_i p_{j-1}$. Due to single-crossing property, these two are sufficient to make sure that consumer i chooses j (verify as an exercise).
- Therefore, consumer i chooses product j iff:

$$\frac{\delta_{j+1} - \delta_j}{p_{j+1} - p_j} < \nu_i < \frac{\delta_j - \delta_{j-1}}{p_j - p_{j-1}}$$

  which implies a set of n cutoff points (see figure).
- Note that, as usual, we normalize the utility from the outside good to be zero for all consumers.

The vertical model (cont.)

- Given a distribution for ν we now have the market share for product j predicted by

$$F\left(\frac{\delta_j - \delta_{j-1}}{p_j - p_{j-1}}\right) - F\left(\frac{\delta_{j+1} - \delta_j}{p_{j+1} - p_j}\right)$$

- Given the distribution and an assumption about the size of the overall market we obtain a one-to-one mapping from the market shares to the δ's, so we can estimate by imposing structures on the δ's and the distribution.
- Note that the vertical model has the property that only prices of adjacent (in terms of prices) products affect the market share, so price elasticity with respect to all other products is zero.
- Is this reasonable? This is a major restriction on the data, and depending on the context you want to think carefully if this is an assumption you want to impose, or that it is too restrictive.

Econometric digression

- So far we assumed that we observe market shares precisely, i.e. that market share data is based on the choice of "infinitely" many consumers.
- This is not always the case (e.g. Berry, Carnall, and Spiller, 1997). In such cases we can get the likelihood of the data to be given by a multinomial distribution of outcomes.
- This gives us

$$L \propto \prod_j s_j(\theta)^{n_j}$$

so that

$$\theta = \arg\max \left[\ln L\right] = \arg\max \left[\sum_j s_j^o \ln s_j(\theta)\right]$$

Econometric digression

- Asymptotically (when $s_j^o = s_j(\theta)$) this is equivalent to

$$\arg\min \left[\sum_j \frac{\left(s_j^o - s_j(\theta)\right)^2}{s_j(\theta)^2} \right]$$

  which is called a minimum χ^2 (or a modified minimum χ^2 when $s_j(\theta)$ is replaced by s_j^o in the denominator).
- This just shows that we should get a better fit on products with smaller market shares. It also shows why we may face more problems when we have tiny market shares.

Logit models

- The basic logit model has
$$U_{ij} = \delta_j + \epsilon_{ij}$$
where $\delta_j = f(X_j, p_j, \xi_j)$ and ϵ_{ij} distributed i.i.d extreme value.
- We get a convenient expression for choice probabilities:
$$\Pr(U_{ij} \geq U_{ik} \forall k) = \frac{\exp(\delta_j)}{1 + \sum_k (\delta_k)}$$
The 1 comes from normalizing the mean utility from the outside good to be zero.
- What are the ϵ_{ij}?
  - unobserved consumer or product characteristics
  - psychological biases (problem with welfare)
  - measurement or approximation errors
- We need it just as we need an ϵ in standard OLS. Without it, the model is unlikely to be able to rationalize the data. (why?)

Logit models (cont.)

- Suppose further that

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

- We can rearrange the market share equation to have $\delta_j = \ln s_j - \ln s_0$, so we have a linear equation we can estimate:

$$\ln s_j - \ln s_0 = X_j\beta - \alpha p_j + \xi_j$$

- The linear form is very useful. We can now instrument for prices using standard IV procedures. This is the main reason people use logit so much: it's "cheap" to do, so you might as well see what it gives you.

Logit models: caveats

- Basic logit model

$$\ln s_j - \ln s_0 = X_j \beta - \alpha p_j + \xi_j$$

- Key drawback: problematic implications for own- and cross-elasticities. To see this, note (and verify at home) that $\frac{\partial s_j}{\partial p_j} = -\alpha s_j (1 - s_j)$ and $\frac{\partial s_j}{\partial p_k} = \alpha s_j s_k$. So:
  - Own-elasticity - $\eta_j = \frac{\partial s_j}{\partial p_j} \frac{p_j}{s_j} = -\alpha p_j (1 - s_j)$ - is increasing in price, which is somewhat unrealistic (we would think people who buy expensive products are less sensitive to price).
  - Cross-elasticity - $\eta_{jk} = \frac{\partial s_j}{\partial p_k} \frac{p_k}{s_j} = \alpha p_k s_k$ - depends only on market shares and prices but not on similarities between goods (think of examples). This is typically called IIA property.

Logit models (cont.)

- Most of the extensions try to correct for the above. Mostly this is not just an issue of the distributional assumption. (What would happen with probit error term?)
- Note that if we just care about ds_j / dx_j and not the elasticity matrix, logit may be good enough. Always remember: whether it is good or not cannot be determined in isolation; it depends on the way it is being used.
- Why do we need ξ_j? this is the analog to the demand-and-supply model, and create the flexibility for us to fit the model. This also shows explicitly the endogeneity of prices, because they are likely to depend on ξ_j and this is why we need to instrument for them (examples).
- Instruments are typically based on the mean independence assumption, i.e. $E(\xi_j|X) = 0$. Does this make sense? What are the assumptions that need to be made to make this go through? Is pre-determination sufficient?

Nested logit

- The basic idea is to relax IIA by grouping the products (somewhat similar idea to AIDS).
- Within each group we have standard logit (with its issues discussed before), but products in different nests have less in common, and therefore are not as good substitutes.
- Formally, utility is given by:

$$U_{ij} = \delta_j + \zeta_{ig}(\sigma) + (1-\sigma)\epsilon_{ij}$$

  with ζ_{ig} being common to all products in group g, and follows a distribution (which depends on σ) that makes $\zeta_{ig}(\sigma) + (1-\sigma)\epsilon_{ij}$ extreme value.
- As σ goes to zero, we are back to the standard logit. As σ goes to one, only the nests matter (so which products do we choose?).

Nested logit, cont.

- A particular nesting, with outside good in one nest and the rest in the other, is relatively cheap to run, so it is used quite often as a robustness check.
- This nesting gives us a linear equation:

$$\ln s_j - \ln s_0 = X_j\beta - \alpha p_j + \sigma \ln(s_{j/g}) + \xi_j$$

  so we can instrument for prices and $s_{j/g}$ and slightly relax the logit assumption.
- One big issue with nested-logit (as with AIDS): need to a-priori classify products. This is not trivial (examples). The following random coefficient models will try to solve this and provide more general treatment (other semi-solution: GEV).

Random coefficients ("BLP")

- Also called mixed logit or heterogeneous logit in other disciplines. These models were around before. The key innovation here is to use these models with aggregate data to obtain a computable estimator with less a-priori restrictions on the substitution pattern.
- Generally, we can write $u_{ij}(X_j, p_j, \xi_j, \nu_i; \theta)$ but we will work with a more specific linear functional form. How restrictive is linearity?. We should ask this question in the context of the economic question we want to answer.
- The model is:

$$U_{ij} = X_j \beta_i - \alpha_i p_j + \xi_j + \epsilon_{ij}$$

  with $\beta_i = \beta + \Sigma \eta_i$ and η_i follows a standardized k-dimensional multi-variate distribution and Σ is a variance-covariance scaling matrix.
- The typical application (e.g. Nevo, 2000) has Σ diagonal and η_i standard normal (but one can make other assumptions, e.g. Berry, Carnall, and Spiller, 1997).

Random coefficients (cont.)

- In either case, with this we can write
$$U_{ij} = \delta_j + \nu_{ij}$$
such that $\delta_j = X_j\beta - \alpha p_j + \xi_j$ and $\nu_{ij} = X_j \Sigma \eta_i + \epsilon_{ij}$.
- Now it is easy to see the difference from the basic logit model: the idiosyncratic error term is not i.i.d but depends on the product characteristics, so consumers who like a certain product are more likely to like similar products.
- How would the substitution matrix look now? Think about the derivatives:
  - $-\alpha s_j(1-s_j)$ becomes $-\int_{\eta_i} \alpha_i s_{ij}(1-s_{ij}) dF(\eta_i)$
  - $\alpha s_j s_k$ becomes $\int_{\eta_i} \alpha_i s_{ij} s_{ik} dF(\eta_i)$
- This achieves exactly what we wanted: substitution which depends on the characteristics (which characteristics?).

Estimating random coefficients

- The key point that facilitates the estimation of this and related models is the inversion, i.e. the possibility to write $\delta(s)$ instead of $s(\delta)$. If this can be done, then we can proceed relatively easy by applying simple GMM restrictions.
- In the previous models, this inversion was carried out analytically. Here that won't work but we can invert numerically, conditional on the "non-linear" parameters of the model, i.e. Σ. Once we have this, we can specify moment conditions. It is important to remember that we need enough moment conditions to identify the Σ parameters as well.

Estimating random coefficients, cont.

- Another problem here is that to compute the integral $s(\delta)$ we need to rely on simulations. The idea: obtain draws from the distribution of η_i and approximate the integral $\int_{\eta_i} s_{ij} dF(\eta_i)$ by $\frac{1}{NS}\sum_{i=1}^{NS} s_{ij}(\eta_i)$. The trade-off here is between more accurate approximation and increased computation time.
- Two computational notes:
  - We take the draws only once, in the beginning, otherwise we never converge.
  - We do not need a whole lot of simulations per market; with many markets the simulation errors average out.

Estimating random coefficients (cont.)

The estimation algorithm (see also Nevo, 2000):

1. Given (δ, Σ) compute $s(\delta, \Sigma)$ using the simulation draws (standard logit per type), as described before.
2. Invert to get $\delta(s, \Sigma)$. This is done numerically by iterating over
$$\delta^{new} = \delta^{old} + (\ln s^o - \ln s(\delta^{old}, \Sigma))$$
Berry shows that this is a contraction (need initial values for δ).
3. Regular GMM of $\delta(s, \Sigma)$ on X, instrumenting for p, and using more moment conditions to identify Σ as well. The search is done numerically, with the added shortcut that the β's enter linearly, so we need to numerically search only over the non-linear parameters.

Note that the formulation has the dimension of β and of Σ the same. This is artificial and not necessary. The former enters the mean utility and the latter enters the substitution pattern. Moreover, the main computational burden is with respect to Σ, so this is where we really want to save on parameters. We can let β be quite rich without much cost.

BLP (1995) Automobiles

- Data on all models marketed 1971 to 1990: annual US sales data, car characteristics, Consumer Reports reliability ratings, miles per gallon.
- Price variable is the list retail price (in $1000s) for the base model, in 1983 dollars.
- Market size is number of households in the US.
- Specifications: simple logit, IV logit, BLP. Price instruments are functions of rival product characteristics and cost shifters.
- Also incorporate a cost model:

$$p = mc + b\left(p, x, \xi; \theta\right)$$

or rewriting with $mc = \exp(w\gamma + \omega)$:

$$\ln\left(p - b\left(p, x, \xi; \theta\right)\right) = w\gamma + \omega.$$

BLP (1995) Automobiles, Results

- Logit model: 1494 of 2217 models have inelastic demands - inconsistent with profit maximization. With IV, allows for unobserved product quality: only 22 models have inelastic demands.
- Full model: most coefficients at least somewhat plausible. Costs: ω accounts for 22% of the estmate variance in log marginal cost. Correlation between ω and ξ is positive (why?).

BLP (1995) Results

TABLE VI

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

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

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

BLP (1995) Results

TABLE VII

SUBSTITUTION TO THE OUTSIDE GOOD

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

BLP (1995) Results

TABLE VIII

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

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

Nevo (2000)

- Ready-to-Eat (RTE) cereal market: highly concentrated, many similar products and yet apparently margins and profits are relatively high. What is the source of market power? Differentiation? Multi-product firms? Collusion?
- Data: market is defined as a city-quarter. IRI data on market shares and prices for each brand-city-quarter: 65 cities, 1Q88-4Q92. Focus on top 25 brands – total share is 43-62%.
- Most of the price variation is cross-brand (88.4%), the remainder is mostly cross-city, and a small amount is cross-quarter.
- Relatively poor "brand characteristics," so model ξ_j as brand "fixed effect" plus market-level "error term". Fixed effect specification differs from random effect set-up in BLP, and is possible because of panel data. Later project brand fixed effect on characteristics.
- Instruments: price of same brand in other city. Identifying assumption: conditional on brand fixed effect, covariation of prices across cities is due to common cost shocks, not demand shocks. (plausible?)

Nevo (2000)

TABLE I

VOLUME MARKET SHARES

	88Q1	88Q4	89Q4	90Q4	91Q4	92Q4
Kellogg	41.39	39.91	38.49	37.86	37.48	33.70
General Mills	22.04	22.30	23.60	23.82	25.33	26.83
Post	11.80	10.30	9.45	10.96	11.37	11.31
Quaker Oats	9.93	9.00	8.29	7.66	7.00	7.40
Ralston	4.86	6.37	7.65	6.60	5.45	5.18
Nabisco	5.32	6.01	4.46	3.75	2.95	3.11
C3	75.23	72.51	71.54	72.64	74.18	71.84
C6	95.34	93.89	91.94	90.65	89.58	87.53
Private Label	3.33	3.75	4.63	6.29	7.13	7.60

Source: IRI Infoscan Data Base, University of Connecticut, Food Marketing Center.

Nevo (2000)

TABLE VI

RESULTS FROM THE FULL MODEL[a]

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

[a] Based on 27,862 observations. Except where noted, parameters are GMM estimates. All regressions include brand and time dummy variables. Asymptotically robust standard errors are given in parentheses.
[b] Estimates from a minimum-distance procedure.

Nevo (2000)

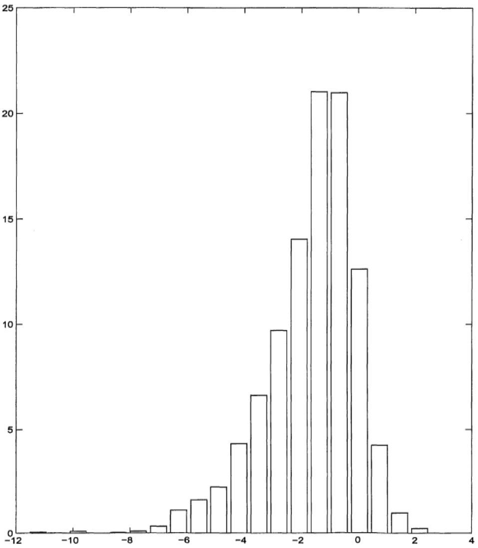

FIGURE 1.—Frequency distribution of taste for sogginess (based on Table VI).

Nevo (2000)

TABLE VIII

MEDIAN MARGINS[a]

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

[a] Margins are defined as $(p - mc)/p$. Presented are medians of the distribution of 27,862 (brand-city-quarter) observations. 95% confidence intervals for these medians are reported in parentheses based on the asymptotic distribution of the estimated demand coefficients. For the Logit model the computation is analytical, while for the full model the computation is based on 1,500 draws from this distribution.

- Compares to accounting PCM as estimated by Cotterill (1996) and concludes that multi-product Bertrand-Nash cannot be rejected.

Consumer Stockpiling

- Demand estimates for CPGs often use time-series variation in prices that comes from sales.
- Problem: short-run and long-run elasticities may be very different if the response to a sale is to "stockpile" inventory at home. Think about something like "cash-for-clunkers" — how much of the sales increase was intertemporal substitution?
- Example: suppose all the toilet paper at the supermarket is marked down 50% for a week, and we observe a 20% increase in demand. This does not mean that if prices were permanently reduced 50% that national consumption of toilet paper would increase 20%!

Consumer Stockpiling: Hendel & Nevo

- Hendel and Nevo (2006, RJE): evidence for stockpiling, e.g. the "post-promotion dip".
- Hendel and Nevo's (2006, EMA): dynamic demand model with consumer inventory as an (unobserved) state variable. Estimate the model using household-level scanner data on laundry detergents. Pretty complicated.
- Hendel and Nevo (2009, WP): a "simpler" method based on a particular model of inventory and sales behavior, that does not require estimation of a complicated dynamic decision proces.

Hendel and Nevo (2006) Results

TABLE VII

LONG-RUN OWN- AND CROSS-PRICE ELASTICITIES[a]

Brand	Size (oz.)	All[b]	Wisk	Surf	Cheer	Tide	Private Label
All[b]	32	0.418	0.129	0.041	0.053	0.131	0.000
	64	0.482	0.093	0.052	0.033	0.085	0.006
	96	0.725	0.092	0.036	0.035	0.100	0.002
	128	−2.536	0.154	0.088	0.059	0.115	0.007
Wisk	32	0.088	0.702	0.046	0.012	0.143	0.006
	64	0.078	0.620	0.045	0.014	0.116	0.004
	96	0.066	0.725	0.051	0.022	0.135	0.009
	128	0.126	−2.916	0.083	0.026	0.147	0.005
Surf	32	0.047	0.061	0.977	0.024	0.369	0.003
	64	0.146	0.086	0.905	0.023	0.158	0.005
	96	0.160	0.101	0.915	0.016	0.214	0.001
	128	0.202	0.149	−3.447	0.039	0.229	0.008
Cheer	64	0.168	0.049	0.027	0.831	0.293	0.001
	96	0.167	0.015	0.008	0.982	0.470	0.001
	128	0.250	0.090	0.058	−4.341	0.456	0.003
Tide	32	0.071	0.085	0.050	0.022	1.007	0.002
	64	0.048	0.055	0.024	0.025	0.924	0.001
	96	0.045	0.063	0.016	0.026	1.086	0.001
	128	0.072	0.093	0.039	0.045	−2.683	0.001
Solo	64	0.066	0.070	0.027	0.021	0.150	0.002
	96	0.219	0.032	0.023	0.033	0.075	0.000
	128	0.127	0.125	0.060	0.043	0.302	0.001
Era	32	0.035	0.155	0.039	0.022	0.425	0.000
	64	0.030	0.103	0.039	0.018	0.304	0.008
	96	0.035	0.168	0.033	0.027	0.352	0.001
	128	0.054	0.192	0.061	0.029	0.513	0.014
Private label	64	0.123	0.119	0.066	0.039	0.081	0.248
	128	0.174	0.266	0.100	0.019	0.072	−2.682
No purchase		0.007	0.002	0.004	0.000	0.013	0.000

[a]Cell entries i and j, where i indexes row and j indexes column, give the percent change in market share of brand i with a 1 percent change in the price of j. All columns are for a 128 oz. product, the most popular size. The results are based on Tables IV–VI.

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

Hendel and Nevo (2006) Results

TABLE VIII

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

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

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

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

Comments and extensions to logit-related models

1. So far we had in mind only aggregate data. How much better can we do with individual-level data?
   1. We can get flexible substitution patterns for free
   2. We may worry less about price endogeneity (why? why do we still need to worry about it?)
   3. With panel dimension, we may be able to identify taste parameters for the unobserved quality

(ref: Goldberg, 1995; "micro BLP", 2004).

Comments and extensions to logit-related models (cont.)

1. Instruments: most use instruments that are based on the exogeneity of the characteristics. As already discussed, this is questionable. It also makes our counterfactuals unlikely to hold for the long run, as characteristics will respond.
   One can use the Hausman-type instruments (similar idea in Nevo, 2001), but they have their issues. Optimally, we would like to have true product-specific cost shifters, but these are hard to find. Once we think about endogenous characteristics, this issue becomes more explicit.

Comments and extensions to logit-related models (cont.)

3. Too many characteristics problem: any new product comes with a new dimension of unobserved tastes (ϵ_{ij}), and a new set of consumers who really like it. Happens even if the new product is identical or inferior to existing products (eg red bus-blue bus).
   - This is likely to bias upwards estimates for markups, and to bias upwards welfare effects of new goods.
   - It does not allow us to use information on goods with zero market shares; the model predicts positive shares.

   One solution: Berry and Pakes, 2002. Like BLP but no ϵ_{ij}. Tricky to recover the mean utility as a function of market shares because: (a) no smooth market share function: they use the vertical model for one coefficient (e.g. price), conditional on the other coefficients; and (b) inversion is not a contraction anymore: they use numerical techniques. Another solution: Bajari and Benkard, 2005. Based on an hedonic approach (and requires a "dense" product space).