Lecture Notes: Estimation of dynamic discrete choice models*

Jean-François Houde
Cornell University

November 7, 2016

*These lectures notes incorporate material from Victor Agguirregabiria's graduate IO slides at the University of Toronto ($http://individual.utoronto.ca/vaguirre/courses/eco2901/teaching_io_toronto.html$).

Demand for storable goods

- **Key references:**
  - Pesendorfer (2002)
  - Erdem, Imai, and Keane (2003)
  - Hendel and Nevo (2006a) and Hendel and Nevo (2006b)
- If consumers can stockpile a storable good, they can benefit from "sales": purchase more than consumption when prices are low.
- This creates a difference between *purchasing* and *consumption* elasticity
  - In the short-run, store-level demand elasticity reflect stockpiling behavior
  - In the long-run, store-level demand elasticity reflects consumption decisions
- Ignoring stockpiling and forward-looking behavior can lead to biases: own and cross price elasticities.
- If consumers differ in the storage costs firms have an incentive to price discriminate: PD theory of sales.
- Ketchup pricing example (Pesendorfer 2002):

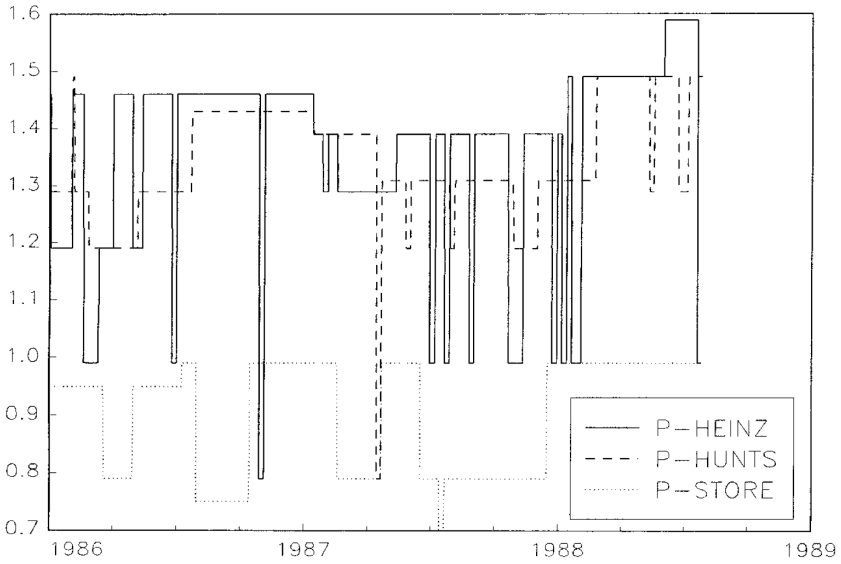


FIG. 1.—Prices for selected products in supermarket 1

- **Hendel and Nevo (2006a):**
  - Estimate an inventory control model *with* product differentiation
  - Two challenges: (i) inventories are unobserved (key state variable), and (ii) differentiation creates a dimensionality problem
  - Propose a simplifying assumption that: (i) reduce the computation burden of the model (reduce the dimension), and (ii) simplify the estimation (three-step algorithm).
- **Data:**
  - Scanner data: Household panel with purchasing decisions from 1991-1993
  - Firms: 9 supermarkets (ignore supermarket choice)
  - Store-level data: choice-set and product attributes (detergent), promotions, and prices

TABLE I

SUMMARY STATISTICS OF HOUSEHOLD-LEVEL DATA[a]

	Mean	Median	Std	Min	Max
Demographics					
Income (000's)	35.4	30.0	21.2	<10	>75
Size of household	2.6	2.0	1.4	1	6
Live in suburb	0.53	—	—	0	1
Purchase of laundry detergents					
Price ($)	4.38	3.89	2.17	0.91	16.59
Size (oz.)	80.8	64	37.8	32	256
Quantity	1.07	1	0.29	1.00	4
Duration (days)	43.7	28	47.3	1	300
Number of brands bought over the 2 years	4.1	3	2.7	1	15
Brand HHI	0.53	0.47	0.28	0.10	1.00
Store visits					
Number of stores visited over the 2 years	2.38	2	1.02	1	5
Store HHI	0.77	0.82	0.21	0.27	1.00

[a]For Demographics, Store visits, Number of brands, and Brand HHI, an observation is a household. For all other statistics, an observation is a purchase instance. Brand HHI is the sum of the square of the volume share of the brands bought by each household. Similarly, Store HHI is the sum of the square of the expenditure share spent in each store by each household.

- Market shares and sales:

TABLE II

BRAND VOLUME SHARES AND FRACTION SOLD ON SALE[a]

	Liquid					Powder				
	Brand	Firm	Share	Cumulative	% on Sale	Brand	Firm	Share	Cumulative	% on Sale
1	Tide	P & G	21.4	21	32.5	Tide	P & G	40	40	25.1
2	All	Unilever	15	36	47.4	Cheer	P & G	14.7	55	9.2
3	Wisk	Unilever	11.5	48	50.2	A & H	C & D	10.5	65	28
4	Solo	P & G	10.1	58	7.2	Dutch	Dial	5.3	70	37.6
5	Purex	Dial	9	67	63.1	Wisk	Unilever	3.7	74	41.2
6	Cheer	P & G	4.6	72	23.6	Oxydol	P & G	3.6	78	59.3
7	A & H	C & D	4.5	76	21.5	Surf	Unilever	3.2	81	11.6
8	Ajax	Colgate	4.4	80	59.4	All	Unilever	2.3	83	
9	Yes	Dow Chemical	4.1	85	33.1	Dreft	P & G	2.2	86	15.2
10	Surf	Unilever	4	89	42.5	Gain	P & G	1.9	87	16.7
11	Era	P & G	3.7	92	40.5	Bold	P & G	1.6	89	1.1
12	Generic	—	0.9	93	0.6	Generic	—	0.7	90	16.6
13	Other	—	0.2	93	0.9	Other	—	0.6	90	19.9

[a]Columns labeled Share are shares of volume (of liquid or powder) sold in our sample, columns labeled Cumulative are the cumulative shares, and columns labeled % on Sale are the percent of the volume, for that brand, sold on sale. A sale is defined as any price at least 5 percent below the modal price, for each UPC in each store. A & H = Arm & Hammer; P & G = Procter and Gamble; C & D = Church and Dwight.

- Descriptive evidence (Hendel and Nevo 2006b)
  - Issue: Inventories are unobserved
  - Duration since previous sale promotion is *positively* correlated with aggregate quantity purchased
  - Proxy for inventory cost (e.g. house size, dog) are negatively correlated with propensity to buy on sale

TABLE 4 **Demand as a Function of Duration from Previous Promotional Activity**

	Laundry Detergents		Yogurt		Soft Drinks	
Variable	Sale	Nonsale	Sale	Nonsale	Sale	Nonsale
Log(price per 16 oz.)	−2.42	−2.40	−1.46	−1.83	−2.59	−1.50
	(.06)	(.02)	(.05)	(.03)	(.05)	(.05)
Duration from previous sale	1.30	.67	1.62	.33	2.99	1.72
	(.58)	(.13)	(.80)	(.24)	(.70)	(.28)
(Duration from previous sale)^2	−1.90	−1.44	6.50	−2.03	−4.13	−3.18
	(1.89)	(.26)	(3.59)	(1.05)	(1.71)	(.40)
Feature	.44	.56	.31	.66	−.01	.56
	(.03)	(.07)	(.02)	(.04)	(.02)	(.04)
Display	1.12	1.19	.67	.96	1.48	1.32
	(.03)	(.02)	(.03)	(.06)	(.02)	(.04)
Duration from previous feature	−.30	−.84	−3.01	−1.46	.55	−.50
	(.23)	(.11)	(.43)	(.21)	(.19)	(.17)
(Duration from previous feature)^2	−.09	1.41	8.48	5.51	−.35	−.69
	(.27)	(.14)	(1.50)	(.72)	(.24)	(.23)
Duration from previous display	−1.21	−.37	−.68	−.29	−2.01	−1.39
	(.19)	(.08)	(.19)	(.09)	(.21)	(.15)
(Duration from previous display)^2	1.05	.04	1.74	.36	1.89	.90
	(.23)	(.12)	(.35)	(.19)	(.29)	(.22)
N	6,681	35,314	9,297	41,226	14,889	22,135

Note: The dependent variable in all regressions is the natural logarithm of quantity purchased (measured in 16-oz. units). Each observation is a brand-size combination in a particular store. Duration from previous sale/feature/display is measured as number of weeks, divided by 100, from previous sale/feature/display for that brand in that store for any size. All regressions include brand and store dummy variables. The regressions in the soft drinks category are for the subsample of cans and include dummy variables for high-demand holiday weeks (July 4, Labor Day, Thanksgiving and Christmas).

TABLE 5 **Correlation Between Household's Fraction of Purchases on Sale and Household Characteristics**

Variable	(1)	(2)	(3)	(4)	(5)	(6)
Constant	.50	.50	.49	.39	.51	.44
	(.02)	(.02)	(.02)	(.02)	(.02)	(.03)
Market 1	−.05	—	−.05	—	—	−.04
	(.01)		(.01)			(.01)
Dog dummy variable	—	.04	.04	—	—	.04
		(.01)	(.01)			(.01)
Cat dummy variable	—	−.001	.005	—	—	.005
		(.016)	(.016)			(.016)
Number of stores	—	—	—	.033	—	.027
				(.006)		(.006)
Average days between shopping	—	—	—	—	−.008	−.005
					(.002)	(.002)
R^2	.045	.037	.051	.059	.042	.080

Note: Dependent variable is the fraction of purchases made during a sale averaged across the three categories: laundry detergents, yogurt, and soft drinks. There are 1,039 observations, where each household is an observation. All regressions also include per-person household income and dummy variables for a male head of household, female working fewer than 35 hours, female working more than 35 hours (excluded category is retired/unemployed), female post-high school education, and head of household is Latino.

TABLE 6 **Differences in Purchasing Patterns Between Sale and Nonsale Purchases**

	Laundry Detergents				Yogurt				Soft Drinks			
		Difference During Sale				Difference During Sale				Difference During Sale		
	Average During		Households		Average During		Households		Average During		Households	
Variable	Nonsale	Total	Within	Between	Nonsale	Total	Within	Between	Nonsale	Total	Within	Between
Quantity	4.79	1.55	1.14	2.22	1.60	.16	.20	.22	5.07	2.79	1.68	4.72
(16 oz.)	(.04)	(.07)	(.07)	(.27)	(.01)	(.02)	(.02)	(.08)	(.07)	(.08)	(.07)	(.42)
Units	1.07	.09	.08	.12	2.63	.99	.80	1.24	4.09	−1.92	−1.19	−1.69
	(.01)	(.01)	(.01)	(.03)	(.03)	(.05)	(.04)	(.16)	(.05)	(.07)	(.06)	(.29)
Size	4.50	.91	.63	1.28	.80	−.19	−.11	−.23	2.93	2.28	1.31	3.84
(16 oz.)	(.03)	(.005)	(.05)	(.20)	(.01)	(.01)	(.01)	(.04)	(.04)	(.05)	(.04)	(.24)
Days from	44.38	6.70	−2.01	29.85	27.35	6.25	−1.27	6.87	24.97	5.21	−2.99	8.24
previous	(.68)	(1.12)	(1.03)	(8.11)	(.59)	(1.10)	(1.03)	(8.85)	(.59)	(.74)	(2.07)	(9.70)
Days to	43.75	8.56	1.95	28.91	26.08	9.87	2.78	21.64	26.57	2.61	2.47	−3.83
next	(.67)	(1.14)	(1.04)	(8.46)	(.59)	(1.09)	(1.03)	(8.53)	(.58)	(.73)	(.64)	(9.83)
Previous	.75	−.29	−.05	−.77	.78	−.31	−.13	−.66	.52	−.25	−.07	−.78
purchase	(.01)	(.01)	(.01)	(.02)	(.01)	(.01)	(.01)	(.03)	(.01)	(.01)	(.01)	(.02)
not on sale												

Note: A sale is defined as any price at least 5% below the modal price of a UPC in a store over the observed period. "Within Households" controls for a household fixed effect, while "Between Households" is the regression of household means. The regressions in the soft drinks category are for the subsample of cans. Standard errors are provided in parentheses.

TABLE 7 **Purchase Conditional on Store Visit and Quantity Purchased Conditional on Purchase by Household as a Function of Price and Promotional Activities**

	Laundry Detergents		Yogurt		Soft Drinks	
Variable	Dependent Variable	Quantity	Dependent Variable	Quantity	Dependent Variable	Quantity
Constant	.08	8.98	.06	.94	.12	4.96
	(.001)	(.64)	(.0005)	(.44)	(.0007)	(1.14)
Inventory/100	−.43	−4.45	−.63	−1.16	−.18	−2.91
	(.01)	(.27)	(.0002)	(.16)	(.006)	(.43)
Price	—	−3.79	—	−.27	—	−5.11
	(.15)		(.08)		(.23)	
Price * sale	—	−1.53	—	.70	—	−3.58
	(.15)		(.09)		(.33)	
Sale	—	1.39	—	−.08	—	1.36
	(.16)		(.12)		(.25)	
Feature	—	.14	—	.11	—	−.06
	(.09)		(.03)		(.13)	
Display	—	.18	—	.14	—	.13
	(.08)		(.04)		(.07)	
N	149,802	12,731	149,802	10,457	149,802	20,995

Note: All results are from linear regressions. The dependent variable in the regressions in columns 1, 3, and 5 is equal to one if the household bought, and zero if the household visited the store and did not buy. In all other regressions the dependent variable is the quantity purchased (measured in 16-oz. units), conditional on purchasing a strictly positive quantity. All regressions also include household, product, and store dummy variables. Prices ($/16 oz.) and promotional variables are for the product purchased. The regressions in the soft drinks category are for the subsample of can purchases. Standard errors are provided in parentheses.

- **Model setup:**
  - J differentiated brands
  - Weekly decisions: consumption c_t, purchases quantity $x_t \in \{0, 1, 2, 3, 4\}$, and brand j_t
  - Utility function:

$$U(c_t, v_t, i_{t+1}) = \gamma \ln(c_t + v_t) - \underbrace{\left[\delta_1 i_{t+1} + \delta_2 i_{t+1}^2\right]}_{\text{Inventory cost}} + m_t$$

    where $m_t = \sum_{j,x} d_{jxt}(\beta a_{jxt} - \alpha p_{jxt} + \xi_{jxt} + \epsilon_{jxt})$ is the indirect utility from purchasing the outside good plus the value of the chosen brand.
  - Inventory transition:

$$i_{t+1} = i_t - c_t + x_t$$

  - **Note:** The function m_t implies that taste for brand only enters the purchasing stage, no the consumption stage. That is, the utility of consumption does not depend on the mix of brands in storage.
  - State variables: $s_t = \{i_t, v_t, \mathbf{p}_t, \mathbf{a_t}, \xi_\mathbf{t}, \epsilon_\mathbf{t}\}$.

- **Two challenges:**
  1. Unobserved state variable: i_t is unobserved to the econometrician.
  2. Curse of dimensionality: $\{\mathbf{p}_t, \mathbf{a_t}, \xi_\mathbf{t}, \epsilon_\mathbf{t}\}$ is $4 \times 5 \times J$ dimension matrix.

Note: In Rust (1987), we also had unobserved state variables (i.e. logit errors). However, we assumed that it was conditionally independent of the observed state and separately additive, which allowed us to work with the Emax function. Not the case here: i_t is serially correlated and non-separable.

- **Model choice-probability:**

$$\Pr(d_{jxt}=1|s_t)=\frac{\exp\left(\beta a_{jxt}-\alpha p_{jxt}+\xi_{jxt}+M(x,s_t)\right)}{\sum_{x',j'}\exp\left(\beta a_{j'x't}-\alpha p_{j'x't}+\xi_{j'x't}+M(x',s_t)\right)}$$

Where,

$$M(x,s_t)=\max_c \gamma\ln(c+v_t)-\left[\delta_1 i_{t+1}+\delta_2 i_{t+1}^2\right]+\beta E\left[V(s_{t+1})|x,c,s_t\right]$$
$$\text{s.t.}\quad i_{t+1}=i_t-c+x$$

- Note that conditional on purchasing size x, the brand choice is a static problem:

$$\begin{aligned}\Pr(d_{jt}=1|s_t,x_t=x) &= \frac{\exp\left(\beta a_{jxt}-\alpha p_{jxt}+\xi_{jxt}+M(x,s_t)\right)}{\sum_{j'}\exp\left(\beta a_{j'xt}-\alpha p_{j'xt}+\xi_{j'xt}+M(x,s_t)\right)}\\ &= \frac{\exp\left(\beta a_{jxt}-\alpha p_{jxt}+\xi_{jxt}\right)}{\sum_{j'}\exp\left(\beta a_{j'xt}-\alpha p_{j'xt}+\xi_{j'xt}\right)}\end{aligned}$$

  This probability can be estimated by MLE: Multinomial logit model.

- This is the **first-step** of the estimation procedure.

- Endogenous prices and advertising? Brand/size fixed-effects, such that $\xi_{jxt}=\xi_{jx}$.

TABLE IV

FIRST STEP: BRAND CHOICE CONDITIONAL ON SIZE[a]

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

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

- How to deal with the dimensionality problem?
- Since the brand-choice is static, we can write down the expected utility (i.e. Emax) conditional on size:

$$\begin{aligned}\omega_t(x) &= E\left(\max_j \beta a_{jxt} - \alpha p_{jxt} + \xi_{jxt} + \epsilon_{jt}\right) \\ &= \ln\left(\sum_{j=1}^{J} \exp\left(\beta a_{jxt} - \alpha p_{jxt} + \xi_{jxt}\right)\right)\end{aligned}$$

- McFadden (1978) calls $\omega_t(x)$ the "inclusive values". It summarizes the information contained in the vector of prices and characteristics available in period t, very much like a quality-adjusted price index.
- Dynamic programming problem re-defined:

$$\begin{aligned}V(i_t, v_t, \omega_t, \epsilon_t) = \max_{c,x} &\gamma \ln(c - v_t) - C(i_{t+1}) + \omega_t(x) + \epsilon_{xt} \\ &+\beta E\left[V(i_{t+1}, v_{t+1}, \omega_{t+1}, \epsilon_{t+1})|i_t, v_t, \omega_t, \epsilon_t, x, c)\right]\end{aligned}$$

- To be able to define the state-vector solely as $(i_t, v_t, \omega_t, \epsilon_t)$, we need to impose an addition assumption on the transition process of ω_t:

$$\Pr(\omega_{t+1}|a_t, p_t, \xi_t) = \Pr(\omega_{t+1}|\omega_t)$$

This assumption as been called the "Inclusive-value sufficiency" assumption (Gowrisankaran and Rysman 2012).

- **Step 2:** Estimate the Markov process for $\omega_t(x)$

TABLE V

SECOND STEP: ESTIMATES OF THE PRICE PROCESS[a]

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

[a]Each column represents the regression of the inclusive value for a size (32, 64, 96, and 128 ounces, respectively) on lagged values of all sizes. The inclusive values were computed using the results in column (x) of Table IV. The four left columns impose the same process for each household type; the four right columns allow for a different process for each type. Reported results are only for households of type 3, that is, households in market 1 with large families. Results for other types are available from the authors.

- **Step 3:** Dynamic decisions
  - Conditional on buying size x, the optimal consumption path solves the following continuous DP problem:

$$v(x|s_t) = \max_{c \geq 0} \quad \gamma \ln(c - v_t) - C(i_t - c + x) + \beta E\left[V(s_{t+1})|s_t, x, c\right]$$

    and the value function takes the familiar form:

$$V(s_t) = \ln\left(\sum_{x=0}^{4} \exp(v(x|s_t))\right)$$

  - The CCP used in the estimation of the dynamic purchasing decision model is:

$$\Pr(x_{it}|s_{it}) = \frac{\exp(v(x_{it}|s_{it}))}{\sum_{x'} \exp(v(x'|s_{it}))}$$

  - Remaining problem: i_t is unobserved.
  - Solution: Integrate out the unobserved initial inventory levels. How?
    - Let $i_{i0} = 0$.
    - Simulate the model for the first T_0 weeks of the data
    - Record inventory level at T_0: i_{T_0} (i.e. initial state variable)
    - Evaluate the likelihood from weeks T_0 to T, as if inventory levels were observed
    - Repeat the process S times, and average the likelihood contribution of individual i.
  - Implicit assumption: Stationary process at time T_0.

TABLE VI

THIRD STEP: ESTIMATES OF DYNAMIC PARAMETERS[a]

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

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

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	64	0.123	0.119	0.066	0.039	0.081	0.248
label	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.

- Two differences between static and dynamic model:
  1. Static model implies larger price coefficient,
  2. and ignores the inventory problem
- Both lead to a larger **own** price elasticity with the static model, than the long-run own price elasticity with the dynamic model.
- Point two leads to a *lower* cross price elasticity (with the static model). Why?
  - In the data the response to sales is mostly coming from people going from not buying to buying the brand on sale.
  - This leads to predict small cross price elasticities with respect to other products, and large cross price elasticity with respect to the outside good.
  - Or more mechanically, in the static model, the choice probability are all relatively small since consumers buy detergent infrequently. The logic cross-partials are equal to the product of the two choice-probabilities: small cross-elasticities.
  - The dynamic model rationalize this phenomenon high unobserved inventories, and conditional choice probabilities (conditional on purchasing). Everybody needs detergent...

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.

Demand for durable goods

Motivation:

- In many settings, the timing of purchase is as important as what you purchase.
- In durable goods markets, this is true because the quality and price of products available at any point in time vary less, than the quality and price of products offer across time.
- Why? Introduction of new products push the technology frontier upwards, and most firms often respond by using dynamic pricing strategy (i.e. discount products that are about to be discontinued).
- Challenges: (i) Micro data on product replacement is rare, and (ii) dynamic "engine" replacement models with product differentiation have very large state space.
- Examples: Gordon (2010), Gowrisankaran and Rysman (2012)

Model: Gowrisankaran and Rysman (2012)

- Indirect utility over product characteristics (omit time):

$$U_{ij} = \beta_i x_j + \xi_j + \varepsilon_{ij} = \delta_j + \mu_{ij} + \varepsilon_{ij} = u_{ij} + \varepsilon_{ij} \tag{1}$$

- Utility over current product:

$$U_{i0} = \beta_i x_0 + \varepsilon_{i0} = u_{i0} + \varepsilon_{i0} \tag{2}$$

  Note: product 0 is either the outside option (not using a camcorder), or the last camcorder purchased. GR assume that products do not depreciate, and there are no sunk replacement costs.

- Bellman equation:
  - Multinomial choice: $a_i \in A_i = \{0, 1, ..., J\}$. Where J is the current number of products available.
  - Common state-space: $s_i = \{u_{i0}, x_1, ..., x_J, p_1, ..., p_J\}$
  - Expected Bellman equation:

$$\begin{aligned} \bar{V}_i(s_i) &= E_\varepsilon \left[\max_{a \in A_i} u_{i,a} - \alpha_i p_{ia} + \varepsilon_{ia} + \beta E_{s'} \left(\bar{V}_i(s') | s_i, a \right) \right] \\ &= E_\varepsilon \left[\max \left\{ v_i(0|s_i) + \varepsilon_{i0}, \max_{a \in A \backslash 0} v_i(a|s_i) + \varepsilon_{ia} \right\} \right] \end{aligned}$$

    where $p_{ia} = 0$ if $a = 0$, and $v_i(a|s_i) = u_{ia} - \alpha_i p_{ia} + \beta E_{s'} \left(\bar{V}_i(s') | s_i, a \right)$.

- **Curse of dimensionality:** (i) The number of grid points to approximate $V_i(s_i)$ grows exponentially with the number of products, and (ii) The calculation of the continuation value involves a $|s|$ dimension integral (i.e. $E_{s'}(V|x)$).

Dimension reduction assumption:

- Recall that if ε_{ij} is distributed according to a type-1 extreme value distribution with location/spread parameters $(0, 1)$, then $\max_{a \in A \setminus 0} v_i(a|x) + \varepsilon_{ia}$ is also EV1 distributed with location parameter:

$$\omega_i(s_i) = \log \left(\sum_{a \in A \setminus 0} \exp(v_i(a|s_i)) \right)$$

  and spread parameter of 1.

- $\omega_i(s_i)$ is the expected discounted value of buying the preferred camcorder available today, instead of holding to the current one (i.e. outside option).

- Therefore, we can rewrite the expected Bellman equation as:

$$\begin{aligned} \bar{V}_i(s_i) &= E_\varepsilon \left[\max \left\{ v_i(0|s_i) + \varepsilon_{i0}, \omega_i(s_i) + \varepsilon_{i1} \right\} \right] \\ &= \log \left(\exp(v_i(0|s_i)) + \exp(\omega_i(s_i)) \right) \end{aligned}$$

- So far, we're simply re-writing the problem so that it looks more like a sequential decision problem: (i) replace or not replace, and (ii) which new product to buy conditional on replacing. The value of replacing is $\omega_i(s_i)$.

- In order to reduce the dimension of the state space, GR introduce a new assumption:
  Inclusive value sufficiency: If $\omega_i(s_i) = \omega_i(\tilde{s}_i)$, then $g_\omega \left(\omega_i(s')|s_i \right) = g_\omega \left(\omega_i(s')|\tilde{s}_i \right)$. Where $g_\omega(x|s_i)$ is the density of the option value ω.

- Importantly, this implies that if two different states have the same option value ω, then they also have the same value function.

- In addition, we assume that ω_i evolves over time according to an AR(1) process:

$$\omega_{i,t+1} = \rho_{i,0} + \rho_{i,1}\omega_{i,t} + \eta_{i,t+1}$$

  where $\eta_{i,t}$ is mean zero shock.

- The IVS assumption, means that consumers can use the scalar ω to forecast the future and evaluate the value relative value of buying now versus waiting.
- New "sufficient" state space: $s_i = (u_{i0}, \omega_i)$.
- New expected value function:

$$\begin{aligned}
\bar{V}_i(u_{i0}, \omega_i) &= E_\varepsilon \left[\max \left\{ u_{i0} + \varepsilon_{i0} + \beta E_{\omega'}[\bar{V}_i(u_{i0}, \omega')|\omega_i], \omega_i + \varepsilon_{i1} \right\} \right] \\
&= \log \left(\exp(u_{i0} + \beta E_{\omega'}[\bar{V}_i(u_{i0}, \omega')|\omega]) + \exp(\omega_i) \right) \\
&= \Gamma_i(u_{i0}, \omega_i | \bar{V}_i)
\end{aligned}$$

  where $\Gamma_i(\bar{V}_i)$ is a contraction mapping from $\bar{V}_i$ to $\bar{V}_i$.

Equilibrium conditions:

1. Rational expectation:

$$\omega_{i,t+1} = \rho_{i,0} + \rho_{i,1}\omega_{i,t} + \eta_{i,t+1}$$

2. Bellman equation:

$$\bar{V}_i(u_{i0}, \omega_i) = \log\left(\exp(u_{i0} + \beta E_{\omega'}[\bar{V}_i(u_{i0}, \omega')|\omega]) + \exp(\omega_i)\right)$$

3. Consistency of the inclusive value:

$$\omega = \log\left(\sum_{a \in A \setminus 0} \exp(u_{i,a} + \beta E_{\omega'}[\bar{V}_i(u_{i,a}, \omega')|\omega]\right)$$

4. Aggregate market share:

$$\hat{s}_{jt} = s_{jt}(\delta, \theta) = \frac{1}{S}\sum_i \sum_{j_0} P_j(u_{i,j_0}, \omega_{it}|\alpha_i, \beta_i) w_{j_0}(\alpha_i, \beta_i)$$

   where $P_j(u_{i,j_0}, \omega_{it}|\beta_i, \alpha_i)$ is the probability of choosing option j conditional on holding option j_0 last period, $w_{j_0}(\alpha_i, \beta_i)$ is the share of consumers of types (α_i, β_i) holding option j_0 in period $t-1$, and S is the number of simulated consumer.

In order to estimate θ, we must minimize the following GMM criterium function subject to the equilibrium conditions:

$$\begin{aligned}
\min_{\theta} \quad & (\xi^T \mathbf{Z})\Omega^{-1}(\xi^T \mathbf{Z})^T \\
s.t. \quad & \delta_{jt} = X_{jt} + \xi_{jt} = s_{jt}^{-1}(\hat{s}, \theta) \\
& \omega_{i,t+1} = \rho_{i,0} + \rho_{i,1}\omega_{i,t} + \eta_{i,t+1} \\
& \bar{V}_i(u_{i0}, \omega_i) = \Gamma_i(u_{i0}, \omega_i|\bar{V}) = \log\left(\exp(u_{i0} + \beta E_{\omega'}[\bar{V}_i(u_{i0}, \omega')|\omega]) + \exp(\omega_i)\right) \\
& \omega_i = \log\left(\sum_{a \in A \setminus 0} \exp(u_{i,a} + \beta E_{\omega'}[\bar{V}_i(u_{i,a}, \omega_i')|\omega_i]\right)
\end{aligned}$$

Solution steps: At every candidate parameter θ

1. Initialization:
   - Sample time-invariant S consumer types: β_i
   - State-space grid: (u_0, ω)
2. *Outer-loop:* Solve $\delta_{jt} = s_{jt}^{-1}(\hat{s}, \theta)$
   (a) Starting values: δ_{jt}^0
   (b) *Inner-loop:* Solve transition parameters $(\rho_{i,0}, \rho_{i,1})$, and Bellman equation $\bar{V}_i(u_0, \omega)$
      i. Initial guesses: $\bar{V}_i^0(u_o, \omega)$ and $(\rho_{i,0}^0, \rho_{i,1}^0)$, for all $i = 1, ..., S$
      ii. Solve inclusive value fixed-point for each t and i:

$$\omega_{i,t} = \log\left(\sum_{a \in A \setminus 0} \exp(u_{i,a} + \beta E_{\omega_{t+1}}[\bar{V}_i^0(u_{i,a}, \omega_{t+1})|\omega_{i,t}]\right)$$

      iii. Estimate AR(1) process (OLS): $\omega_{i,t+1} = \rho_{i,0}^1 + \rho_{i,1}^1 \omega_{i,t} + \eta_{i,t+1}$
      iv. Update value function:

$$\bar{V}_i^1(u_0, \omega) = \log\left(\exp(u_0 + \beta E_{\omega'}[\bar{V}_i^0(u_0, \omega')|\omega]) + \exp(\omega)\right)$$

      v. Convergence check: $||\bar{V}^1 - \bar{V}^0|| < \epsilon_1$, $||\rho_{i,0}^1 - \rho_{i,0}^0|| < \epsilon_2$ and $||\rho_{i,1}^1 - \rho_{i,1}^0|| < \epsilon_3$
   (c) Calculate predicted market shares:

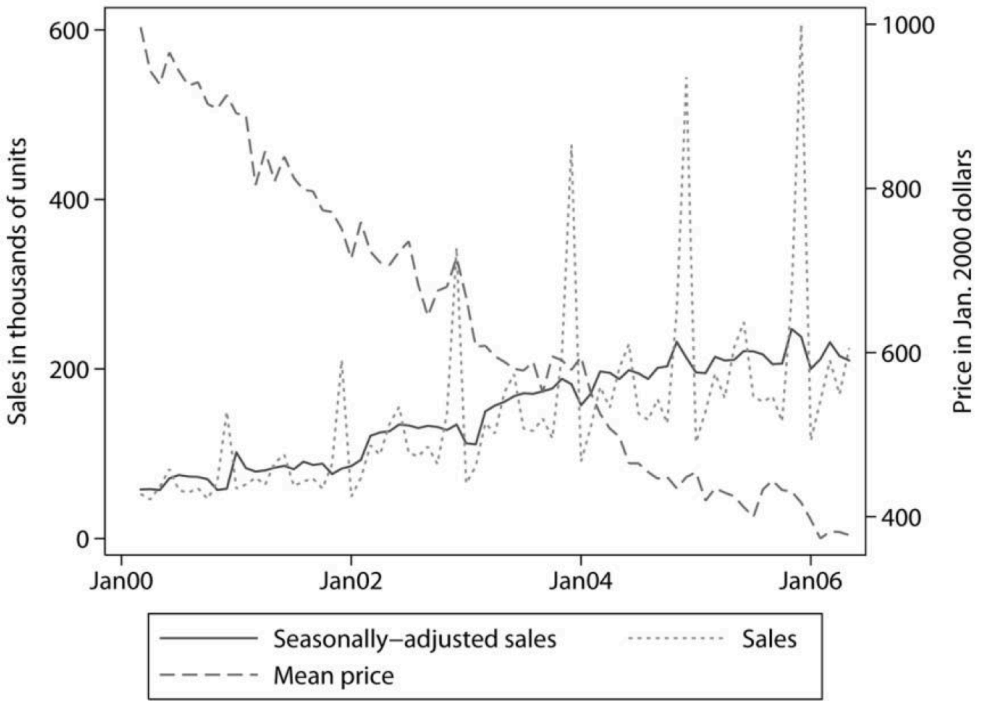
$$\hat{s}_{jt} = s_{jt}(\delta^0, \theta) = \frac{1}{S} \sum_i \sum_{j_0} P_j(u_{i,j_0}, \omega_{it}|\alpha_i, \beta_i) w_{j_0}(\alpha_i, \beta_i)$$

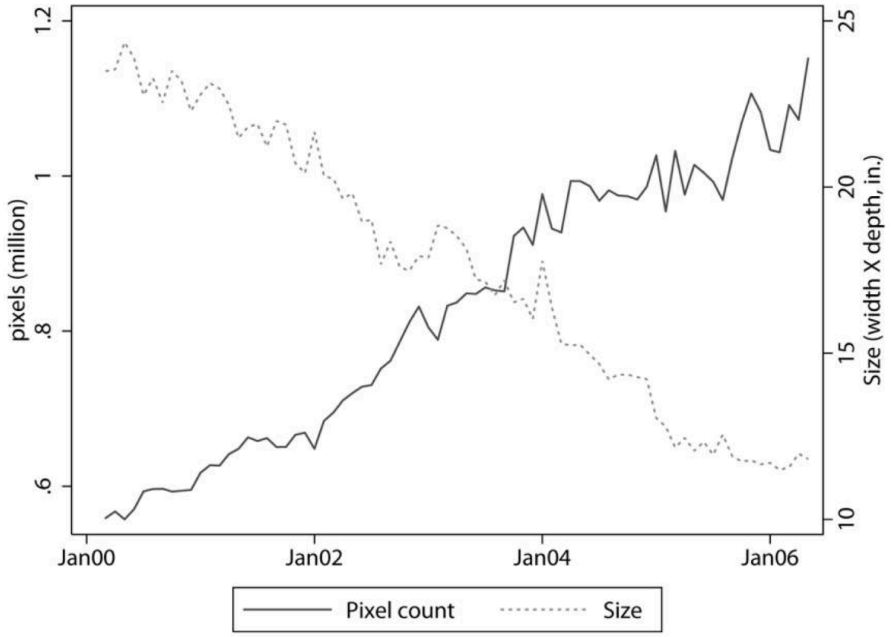
   where $P_j(u_{i,j_0}, \omega_{it}|\alpha_i, \beta_i) = \exp(v_0(u_{i,j_0}))/\left(\exp(\omega_{i,t}) + \exp(v_0(u_{i,j_0}))\right)$ if $j = j_0$, and

$$P_j(u_{i,j_0}, \omega_{it}|\alpha_i, \beta_i) = \frac{\exp(\omega_{i,t})}{\exp(\omega_{i,t}) + \exp(v_0(u_{i,j_0}))} \times \frac{\exp(v_{j,t}(u_{jt}, \omega_t))}{\exp(\omega_{i,t})}$$

Applications: Video-camcorder market between 2000 and 2006

- **Main data-set:** Market shares and characteristics of 383 models and 11 brands, from March 2000 to May 2006
- **Auxiliary data:** Aggregate penetration and new-sales rates by years





Estimation results

Parameter	Base Dynamic Model (1)	Dynamic Model without Repurchases (2)	Static Model (3)	Dynamic Model with Micro Moment (4)
Mean coefficients (α):				
Constant	−.092 (.029)*	−.093 (7.24)	−6.86 (358)	−.367 (.065)*
Log price	−3.30 (1.03)*	−.543 (3.09)	−.099 (148)	−3.43 (.225)*
Log size	−.007 (.001)*	−.002 (.116)	−.159 (.051)*	−.021 (.003)*
(Log pixel)/10	.010 (.003)*	−.002 (.441)	−.329 (.053)*	.027 (.003)*
Log zoom	.005 (.002)*	.006 (.104)	.608 (.075)*	.018 (.004)*
Log LCD size	.003 (.002)*	.000 (.141)	−.073 (.093)	.004 (.005)
Media:				
DVD	.033 (.006)*	.004 (1.16)	.074 (.332)	.060 (.019)*
Tape	.012 (.005)*	−.005 (.683)	−.667 (.318)*	.015 (.018)
Hard drive	.036 (.009)*	−.002 (1.55)	−.647 (.420)	.057 (.022)*
Lamp	.005 (.002)*	−.001 (.229)	−.219 (.061)*	.002 (.003)
Night shot	.003 (.001)*	.004 (.074)	.430 (.060)*	.015 (.004)*
Photo capable	−.007 (.002)*	−.002 (.143)	−.171 (.173)	−.010 (.006)
Standard deviation coefficients ($\Sigma^{1/2}$):				
Constant	.079 (.021)*	.038 (1.06)	.001 (1,147)	.087 (.038)*
Log price	.345 (.115)*	.001 (1.94)	−.001 (427)	.820 (.084)*

NOTE.—Standard errors are in parentheses. All models include brand dummies, with Sony excluded. There are 4,436 observations.

* Statistically significant at the 5 percent level.

Implied elasticities:

- A market-wise *temporary* price increase of 1% leads to:
  - A contemporaneous decrease in sales of 2.55 percent
  - Most of this decrease is due to delayed purchases: 44 percent of the decrease in sales is recaptured over the following 12 months.
- The same market-wise *permanent* price increase leads to a 1.23 percent decrease in demand (permanent).
- The difference is more modest when we consider a product-level price increase: 2.59 percent versus 2.41 percent (for the leading product in 2003).
- Why? Consumers substitute to competing brands when the change is permanent in about the same magnitude as the delayed response when the price change is temporary.

References

Erdem, T., S. Imai, and M. P. Keane (2003). Brand and quantity choice dynamics under price uncertainty. *Quantitative Marketing and Economics 1*, 5–64.

Gordon, B. (2010, September). A dynamic model of consumer replacement cycles in the pc processor industry. *Marketing Science 28*(5).

Gowrisankaran, G. and M. Rysman (2012). Dynamics of consumer demand for new durable goods. *Journal of Political Economy 120*, 1173–1219.

Hendel, I. and A. Nevo (2006a). Measuring the implications of sales and consumer stockpiling behavior. *Econometrica 74*(6), 1637–1673.

Hendel, I. and A. Nevo (2006b, Fall). Sales and consumer inventory. *Rand Journal of Economics*.

McFadden, D. (1978). *Spatial Interaction theory and residential location*, Chapter Modelling the choice of residential location, pp. 531–552. North Holland and Co.

Pesendorfer, M. (2002). Retail sales: A study of pricing behavior in supermarkets*. *Journal of Business 75*(1), 33–66.

Rust, J. (1987). Optimal replacement of gmc bus engines: An empirical model of harold zurcher. *Econometrica: Journal of the Econometric Society 55*(5), 999–1033.