Dynamic Demand

Aviv Nevo

University of Pennsylvania and NBER

Spring 2017

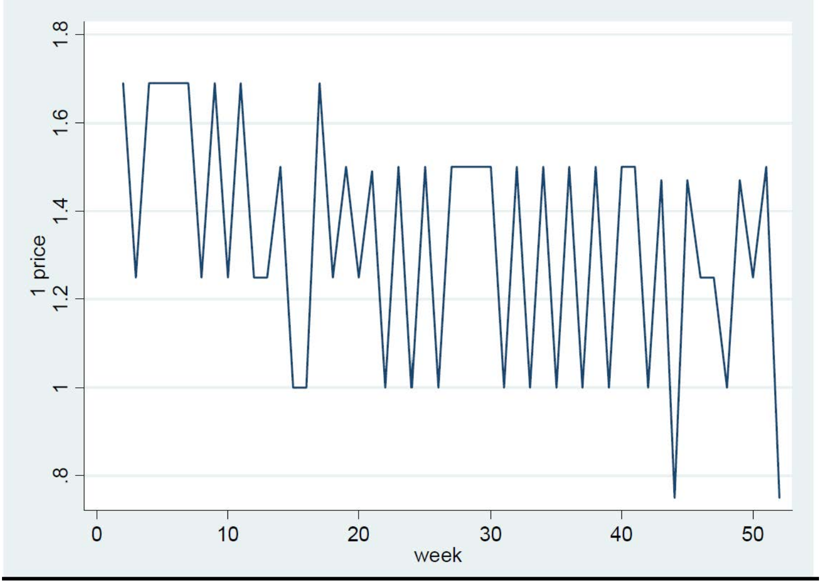
Introduction

- Up to now we focused on static demand
- Dynamics can arise for different reasons. e.g,
  - storable products
  - durable products
  - habit formation
  - learning
  - switching costs
- I will focus on storable and durable products and discuss
  - the biases in static estimation (if dynamics are present)
  - models of estimating dynamics
  - estimation with consumer and aggregate data
- I will focus on modeling and not the details of estimation.

Challenges in estimating demand for Diff Prod

- As we saw DC is a solution to the "too many products" problem
- For dynamics:
  - need to keep track of attributes and prices of all products
  - interacted with consumer attributes (since different consumers care about different goods differently)
  - not practical, hence, need to reduce the dimension
- We will see two different solutions

Storable Products: A typical pricing pattern



Why are there sales?

1. A change in (static) cost, demand or retailer inventory cost
2. Search and mixed strategies (Varian, 1980; Salop and Stiglitz, 1982)
3. Retailer behavior: multi-category pricing
4. Intertemporal price discrimination (Sobel, 1984; Conlisk Gerstner and Sobel, 1984; Pesendorfer, 2002; Narasimhan and Jeuland, 1985; Hong, McAfee and Nayyar, 2002)

What do consumers do? They store

- Pesendorfer (2002): aggregate demand depends on duration from previous sale
- Hendel and Nevo (2006a): demand accumulation and demand anticipation effects
  - HH frequency of purchases on sales correlated with proxies of storage costs
  - when purchasing on sale, longer duration to next purchase (within and across HH)
  - proxies for inventory is (i) negatively correlated with quantity purchased and (ii) the probability of purchasing
- Extensive Marketing Literature: Post-Promotion Dip Puzzle (Blattberg and Neslin, 1990)

Just in case you doubt ...

Table 1: Quantity of 2-Liter Bottles of Coke Sold			
	$S_{t-1} = 0$	$S_{t-1} = 1$	
$S_t = 0$	247.8 (4.8)	199.4 (5.5)	227.0 (3.6)
$S_t = 1$	763.4 (5.6)	531.9 (4.5)	622.6 (3.5)
	465.0 (4.0)	398.9 (3.8)	

Implications for Demand Estimation

- When consumers can store purchases and consumption differ
  - purchases driven by consumption and storage
  - object of interest are preferences
- 2 separate issues with static demand estimation
- Econometric bias: omitted variable that might be correlated with price
- Difference between SR and LR response
- For most applications we care about LR response

Static Estimates vs LR responses

- Static estimation overstates own price effects
  - purchase response to a sale mis- attributed to price sensitivity
  - purchase decrease after a sale mis-attributed to price sensitivity
  - the demand response to a permanent change in price is likely much smaller
- Underestimate cross price effects
  - because of storage "effective" and current prices differ
  - consider the period after a sale of a competing product
    - "effective" (cross) price is the sale period purchase price
    - observed price is higher, but not accompanied by a decline in purchases
    - hence mis-attributed to low cross price sensitivity

Model of consumer stockpiling

The per period utility consumer i obtains from consuming in t

$$u_i(\overrightarrow{c_t}, \overrightarrow{\nu_t}) + \alpha_i m_t$$

choose consumption, brand and size to

$$\max \sum_{t=1}^{\infty} \delta^{t-1} \mathbb{E}[u_i(\overrightarrow{c_t}, \overrightarrow{\nu_t}) - C_i(\overrightarrow{i_t}) + a_{jxt}\beta_i - \alpha_i p_{jxt} + \xi_{jxt} + \varepsilon_{ijxt} | \; s_1]$$

$$\begin{aligned} s.t. \quad 0 &\leq \overrightarrow{i_t}, \quad 0 \leq \overrightarrow{c_t}, \quad \sum_{j,x} d_{jxt} = 1, \\ i_{j,t+1} &= i_{j,t} + \sum_x d_{jxt} x_t - c_{j,t} \quad j = 1, ..., J \end{aligned}$$

Model (cont)

- Stochastic structure
  - ε_{jxt} is i.i.d. extreme value type 1
  - ν_t is i.i.d. over time and across consumers
  - prices (and advertising) follow a first order Markov process.
- The first two can be relaxed at a significant computational cost
- First order price process can be somewhat relaxed (will see below)
- We will see how we deal with price endogeneity

The Value Function

Value function:

$$V_i(\overrightarrow{i_t}, \overrightarrow{p_t}, \overrightarrow{\nu_t}, \overrightarrow{\varepsilon_t}) =$$

$$\max_{\overrightarrow{c},j,x} \left\{ \begin{array}{l} u_i(\overrightarrow{c}, \overrightarrow{\nu_t}) - C_i(\overrightarrow{i_t}) + a_{jxt}\beta_i - \alpha_i p_{jxt} + \xi_{jxt} + \varepsilon_{ijxt} \\ +\delta \ \mathbb{E}\left[V_i(\overrightarrow{i_{t+1}}, \overrightarrow{p_{t+1}}, \overrightarrow{\nu_{t+1}}, \overrightarrow{\varepsilon_{t+1}}) \mid \overrightarrow{i_t}, \overrightarrow{p_t}, \overrightarrow{\nu_t}, \overrightarrow{\varepsilon_t}, \overrightarrow{c}, j, x \right] \end{array} \right\}$$

The integrated value function:

$$EV_i(\overrightarrow{i_t}, \overrightarrow{p_t}) =$$

$$\max_{\overrightarrow{c},j,x} \int \ln \left(\sum_{j,x} \exp \left\{ \begin{array}{l} u_i(\overrightarrow{c}, \overrightarrow{\nu_t}) - C_i(\overrightarrow{i_t}) + a_{jxt}\beta_i - \alpha_i p_{jxt} + \xi_{jxt} \\ +\delta \ \mathbb{E}\left[EV_i(\overrightarrow{i_{t+1}}, \overrightarrow{p_{t+1}}) \mid \overrightarrow{i_t}, \overrightarrow{p_t}, \overrightarrow{c}, j, x \right] \end{array} \right\} \right)$$

- Note the dimension of the state space

Reducing the State Space: Holdings

Assumption A1:

$$U_i(\overrightarrow{c}_t, \overrightarrow{\nu_t}) = U_i(c_t, \nu_t) \text{ and } C_i(\overrightarrow{i_t}) = C_i(i_t)$$

where

$$c_t = \mathbf{1}'\overrightarrow{c_t}, \ \nu_t = \mathbf{1}'\overrightarrow{\nu_t}, \text{ and } i_t = \mathbf{1}'\overrightarrow{i_t}.$$

- In words, no differentiation in usage (at least not in NL part)
  - differentiation in purchase NOT consumption
  - how should we think of this?
  - could relax somewhat by thinking of "segments"

Therefore

$$EV_i(\overrightarrow{i_t}, \overrightarrow{p_t}) = EV_i(i_t, \overrightarrow{p_t})$$

Reducing the State Space: Holdings

- Can further show: $c_k^*(s_t; x, k) = c_j^*(s_t; x, j) = c^*(s_t; x)$.
- In words, consumption does not depend on brand purchased.
- Therefore,

$$EV_i(i_t, \overrightarrow{p_t}) =$$
$$\max_{c,x} \int \ln \left(\sum_x \exp \left\{ \begin{array}{c} u_i(c, v_t) - C_i(i_t) + \omega_{ixt} + \\ +\delta \; E\left[EV_i(i_{t+1}, \overrightarrow{p}_{t+1}) \mid i_t, \overrightarrow{p_t}, c, x \right] \end{array} \right\} \right) dF_\nu(v_t)$$

- In words, the dynamic problem can be seen as a problem of choosing size

Reducing the State Space: Prices

- A key concept, the **Inclusive Value** or (social surplus)
- Assume ε_{ijt} are distributed i.i.d. extreme value,
- The inclusive value from a subset A of alternatives is:

$$\omega_{iAt} = \ln\left(\sum_{j\in A} \exp\left\{a_{jt}\ \beta_i - \alpha_i\ p_{jt} + \xi_{jt}\right\}\right)$$

- Note: ω_{iAt} is individual specific
- A natural way to reduce the set space
- Simplifies the problem by showing/assuming that
  - expected flow utility depends only on this statistic
  - the transition can be summarized with this statistic

Reducing the State Space: Prices

Assumption A2: $F(\overrightarrow{\omega}_{i,t+1} \mid \overrightarrow{p_t}) = F(\overrightarrow{\omega}_{i,t+1} \mid \overrightarrow{\omega}_{it}(\overrightarrow{p_t}))$
where $\overrightarrow{\omega}_{it}$ is a vector with the IV for each size.

- $\overrightarrow{\omega}_{it}$ contains all the information needed to compute the transition probabilities
- A strong assumption that can be somewhat relaxed (and tested), jointly with first-order Markov
- Now:

$$EV_i(i_t, \overrightarrow{p_t}) = EV_i(i_t, \overrightarrow{\omega}_{it}(\overrightarrow{p_t}))$$

- Note the reduction in the dimension of the state space

Data and Identification

- Data
  - consumer level: history of purchase
- Identification
  - no formal proof
  - informally: parameters identified from time series of purchases
    - ex: holding total quantity fixed, average duration between purchases determines storage costs
- Price endogeneity:
  - assume $\xi_{jxt} = \xi_{jx}$ (or $x\xi_j$) and include FE (and feature/display)
  - could nest BLP inversion

Estimation

- Follow the "nested algorithm" approach (Rust, 87)
    - guess a value of the parameters
    - solve the DP
    - use the solution to compute likelihood of data
    - repeat until likelihood is max.
- Use the model to simulate the optimal unobserved policy (consumption) and state (inventory)
    - the nested algorithm provides a natural way to do this,
    - could also use the EM algorithm of Arcidiacono and Miller (2008);

Splitting the Likelihood

- Enrich the model (and speed up computation) by splitting the estimation into $\mathbb{P}(j|x)$ and $\mathbb{P}(x)$.
- 3 step estimation:
  - (static) conditional Logit of brand given size
  - use the estimates to compute $\overrightarrow{\omega}_{it}$ and estimate transitions
  - estimate the dynamic choice problem: purchase/no purchase and size

Splitting the Likelihood

- The split if the likelihood follows from the above assumptions, plus

Assumption A3 (conditional independence of heterogeneity):

$$F(\alpha_i\ , \beta_i | x_t, \overrightarrow{p_t}, D_i) = F(\alpha_i\ , \beta_i | \overrightarrow{p_t}, D_i)$$

- Restricts unobserved heterogeneity in brand choice
  - the choice of size does not tell anything about the distribution of heterogeneity
- Can allow for rich demographics and consumer FE
- Trade-off between speed and richer model vs unobserved heterogeneity
  - richer model in additional variables and observed heterogeneity
- Can estimate the model even without this assumption

Results

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									✓	✓

128 oz.					
All[b]	Wisk	Surf	Cheer	Tide	Private Label
0.14	0.17	0.17	0.18	0.21	0.34
1.23	0.09	0.11	0.09	0.15	0.22
0.16	0.22	0.14	0.22	0.25	0.20
0.08	1.42	0.08	0.13	0.18	0.11
0.18	0.18	0.12	0.18	0.22	0.28
0.12	0.11	1.20	0.08	0.15	0.14
0.14	0.24	0.16	0.14	0.22	0.24
0.09	0.12	0.06	0.89	0.15	0.07
0.22	0.28	0.16	0.26	0.22	0.37
0.11	0.16	0.08	0.13	1.44	0.31
0.17	0.15	0.15	0.30	0.30	0.28
0.07	0.07	0.06	0.16	0.17	0.21
0.43	0.17	0.15	0.22	0.19	0.35
0.19	0.08	0.09	0.11	0.10	0.22
0.32	0.22	0.15	0.26	0.31	0.25
0.16	0.12	0.13	0.10	0.27	1.29
1.80	7.60	2.26	14.11	2.38	10.86

Motivation for simple demand model

- Previous results suggest that neglecting demand dynamics may lead to inconsistent estimates
- Yet
  - the estimation was quite complex (even after simplifications)
  - requires consumer level panel (not always available)
  - difficult to derive supply model
  - especially true if demand is an input into an equilibrium model
- I will now discuss a simple model of demand anticipation (based on Hendel and Nevo, 2011)
- The model:
  - easy to estimate with aggregate data
    - key: storage technology (periods to store, not physical units)
  - makes supply tractable

Simple Model Outline

- Assumptions
  - A1(hetero) 2 consumer types: proportion ω don't store
  - A2(storage) inventory is free, lasts for T periods
  - A3 (expectations) perfect foresight of future prices
  - A3': rational expectations
- Storers and non-storers may have different preferences:

$$U_t^S(\mathbf{q}, m) = u_t^S(\mathbf{q}) + m \text{ and } U_t^{NS}(\mathbf{q}, m) = u_t^{NS}(\mathbf{q}) + m$$

$q = [q_1, q_2, ..., q_N]$ and m is the outside good

- Absent storage: $\mathbf{q}_t^S = Q_t^S(\mathbf{p}_t)$ and $\mathbf{q}_t^{NS} = Q_t^{NS}(\mathbf{p}_t)$
- Denote purchases by X and consumption by Q

Purchasing Patterns

- Under A1-A3 and $T = 1$
  - four states defined by sale/no-sale previous and current period
  - SS, NN, NS and NS
  - NS = sale today and non-sale last period (week)
- Purchases by storers

$$X^S_{jt} = \begin{cases} Q^S_{jt}(p_{jt}, p^{eff}_{-jt}) \\ 0 \\ Q^S_{jt}(p_{jt}, p^{eff}_{-jt}) \\ 0 \end{cases} + \begin{matrix} 0 \\ 0 \\ Q^S_{jt+1}(p_{jt}, p^{eff}_{-jt+1}) \\ Q^S_{jt+1}(p_{jt}, p^{eff}_{-jt+1}) \end{matrix} \quad \text{in} \quad \begin{matrix} NN \\ SN \\ NS \\ SS \end{matrix}$$

- effective price: $p^{eff}_{jt} = \min\{p_{jt-1,} p_{jt}\}$
- Non-storers always contribute $Q^{NS}_j(p_{jt}, p_{-jt})$

Comments on the Model's Assumptions

- A1: in principle decision of whether to be a "storer" should be endogenous
  - A1 can be seen as an assumption on the storage cost distribution
  - fixed proportion, ω, can be made $\omega(p)$
- A2: Storage technology
  - allows us to simplify the state space: there are no left overs to carry as a state variable
  - it detaches the storage decision of different products: the link –between products– is captured by effective prices
  - taken literally, fits perishable products
  - Extensions:
    - easy to allow for $T > 1$
    - heterogeneity across consumers in T
- PF easier to work with but RE doable

How Do We Recover Preferences?

- For simplicity, assume 1 product, $T = 1$, and $Q_t^S(p_t) = Q^S(p_t) + \varepsilon_t$ (where ε_t is an iid error)

$$X_t = Q_t^{NS}(p_t) + \begin{cases} Q_t^S(p_t) \\ 0 \\ Q_t^S(p_t) + Q_{t+1}^S(p_t) \\ Q_{t+1}^S(p_t) \end{cases} \quad in \quad \begin{matrix} NN \\ SN \\ NS \\ SS \end{matrix}$$

- SN periods identify Q^{NS} (for "non-sale" prices)
- From SN and NN we can identify $Q^S(p_t) = X^{NN} - X^{SN}$
- From NS and SS we can identify $Q^S(p_t) = X^{NS} - X^{SS}$ and $Q^{NS}(p_t) = 2X^{SS} - X^{NS}$
- The model is NP identified
    - over-identified with parametric assumptions or overlap in prices (depends on how "sale" is defined)
- Same idea with more products and $T > 1$

Estimation

- Estimate the model by minimizing the distance between observed and predicted purchases
- The estimation controls for prices (own and competition), can control for other factors, and account for endogenous prices
- Can also use GMM/IV
- Need to account for store fixed effects
- Can enrich model to allow for $T > 1$, more types, heterogeneity and more flexible demand systems;

An Empirical Example: Demand for Colas

- Data: Store-level scanner data
  - weekly observations
  - 8 chains in 729 stores in North East
  - Due to data problems will only use 5 chains
  - focus on 2 liter bottles of Coke, Pepsi and Store brands
- We estimate linear demand, allowing for store fixed effects
- A sale is defined as a price $\leq \$1$
- $$\log q^h_{jst} = \omega^h \alpha_{sj} - \beta^h_j p_{jst} + \gamma^h_{ji} p_{ist} + \varepsilon_{jst}, \quad j = 1,2 \quad i = 3-j \quad h = \dots$$

Demand Estimates: T=1

Table 3: Estimates of the Demand Function

	Static Model		Dynamic Models					
			PF		PF-Alt Sale Def		RE	
	Coke	Pepsi	Coke	Pepsi	Coke	Pepsi	Coke	Pepsi
P_{own} non-storers	-2.30	-2.91	-1.41	-2.11	-1.49	-2.12	-1.27	-1.98
	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)
P_{cross} non-storers	0.49	0.72	0.61		0.71		0.63	
	(0.01)	(0.01)	(0.01)		(0.01)		(0.01)	
P_{own} storers			-4.37	-5.27	-4.38	-5.08	-5.57	-6.43
			(0.06)	(0.06)	(0.07)	(0.07)	(0.12)	(0.12)
P_{cross} storers			0.61		0.34		2.12	
			(0.04)		(0.04)		(0.09)	
ω	–	–	0.14		0.10		0.21	
			(0.01)		(0.01)		(0.02)	
# of observations			45434		45434		30725	

Durable Products

- A typical pricing pattern: prices declining over time
- Implications for demand estimation
  - with no repeat purchase, need to
    - account for change in distribution of consumers
    - option value of waiting
  - with repeat purchase, need to
    - account for variation in outside good
    - account for expectations

Example of implications

- $WTP \sim U[0, 1]$, total mass of 100
- Consumers are myopic: buy if $p_t \leq WTP$
- No repeat purchase
- Aggregate demand: $Q_t = 100 - 100P_t$
- Suppose we observe a seq of prices $(0.9, 0.8, 0.7, ..., 0.1)$.
- Therefore, $q_t = 10$
- Static estimation will find no price sensitivity (i.e., underestimates the price sensitivity)
- Depends on the distribution of WTP and prices
- Even more complicated to sign with repeat purchases

Model

$$u_{ijt} = \omega^f_{ijt} - \alpha_i p_{jt} + \varepsilon_{ijt}$$

where the flow utility is

$$\omega^f_{ijt} = a_{jt}\beta_i + \xi_{jt}.$$

If the consumer does not purchase she gets the utility

$$u_{i0t} = \omega^f_{i0t} + \varepsilon_{i0t}$$

where

$$\omega^f_{i0t} = \begin{cases} 0 & \text{if no previous purchase} \\ \omega^f_{i\widehat{j}\widehat{t}} & \text{if last purchase was product } \widehat{j} \text{ at time } \widehat{t} \end{cases}.$$

Value Function

with repeat purchase

$$V_i(\varepsilon_{it}, \omega^f_{i0t}, \Omega_t) = \max_{j=0,\dots J} \left\{ u_{ijt} + \delta \mathbb{E}[EV_i(\omega^f_{ijt}, \Omega_{t+1}|\Omega_t] \right\}$$

with no repeat purchase

$$V_i(\varepsilon_{it}, \Omega_t) = \max \left\{ \varepsilon_{i0t} + \delta \ \mathbb{E}[EV_i(\Omega_{t+1}|\Omega_t] \ , \ \max_{j=1,\dots,J} u_{ijt} \right\}$$

Reducing the state space: holdings

- Parallel to stockpiling problem
- Single holding/no interaction in utility
- Quantity of holding versus quality of holding

Reducing the state space: prices/quality

$$\omega_{ijt}^{D} = \omega_{ijt}^{f} - \alpha_i p_{jt} + \delta \mathbb{E}[EV_i(\omega_{ijt}^{f}, \Omega_{t+1})|\Omega_t]$$

the *dynamic inclusive value:*

$$\omega_{it}^{D}(\Omega_t) = \ln\left(\sum_{j=1}^{J} \exp(\omega_{ijt}^{D})\right).$$

- Static IV provides a summary of (exogenous) prices and attributes of available products.
- Dynamic IV also includes (endogenous) future behavior of the agent.
- With repeat purchases need the latter (w/o repeat purchase can do with static IV)

Assumption A2′: $F(\omega_{i,t+1}^{D} \mid \Omega_t) = F(\omega_{i,t+1}^{D} \mid \omega_{it}^{D}(\Omega_t))$

Therefore: $EV_i(\omega_{i0t}^{f}, \Omega_t) = EV_i(\omega_{i0t}^{f}, \omega_{it}^{D})$

Econometrics

- Data
  - can rely on consumer level data
  - typically use aggregate/market level data
- Identification
  - lines up WTP in the CS with variation over time
- Estimation
  - Nests the solution of the DP within the BLP estimation algorithm

Gowrisankaran and Rysman (2011)

- Study demand for camcorders
- Compare dynamic demand estimates to static ones

Table 1: Parameter estimates

Parameter	Base dynamic model	Dynamic model without repurchases	Static model	Dynamic model with micro-moment
	(1)	(2)	(3)	(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) *
Media: tape	.012 (.005) *	-.005 (.683)	-.667 (.318) *	.015 (.018)
Media: HD	.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 (1147)	.087 (.038) *
Log price	.345 (.115) *	.001 (1.94)	-.001 (427)	.820 (.084) *