Dynamic Demand I:
Durable Goods + Storable Goods

C.Conlon

Grad IO

November 10, 2017

Dynamic Demand

- Earlier this term we looked at *static models of product differentiation* such as BLP (1995).
- We have also looked at single agent models of dynamic behavior such as Rust (1987).
- What if we could put those two together? Why?

Two Examples

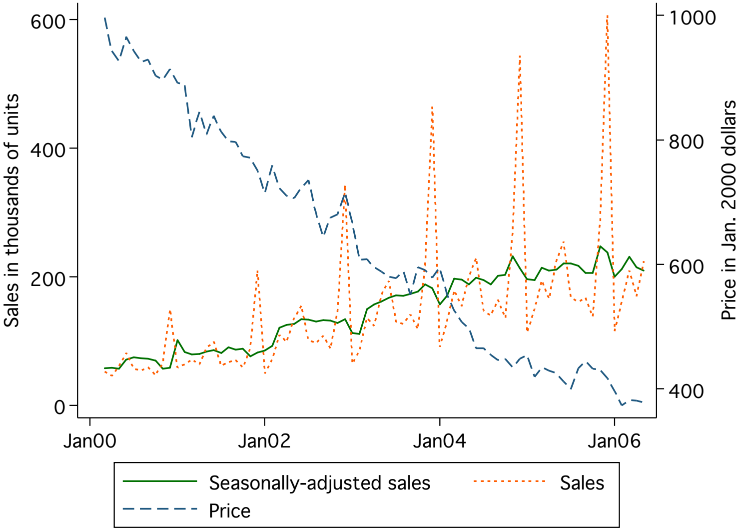
- Durable Goods
  - We are often interested in the case where a firms' products compete not only against products of other firms, but also with their own products over time. (Coase 1972).
  - Really this is a feature of any kind of durable (cars, washing machines, etc.).
  - In high-tech products, we have 'S'-shaped penetration curve. At some point we cut the prices of (computers, cameras, TV's) and the sales go down rather than up. Why?
- Storable Goods
  - Laundry detergent goes on discount: ≈ 1 week in every 4-5.
  - A large fraction of overall quantity is sold during these discount weeks.
  - This would imply highly elastic demand. (1) Laundry detergent business must be extremely competitive. (2) There would be a substantial response to a permanent price change.
  - Seems unlikely that (1) and (2) are true.

Dynamic Demand (Gowrisankaran Rysman)

Figure 1: Average non-indicator characteristics over time

Dynamic Demand (Gowrisankaran Rysman)

Figure 3: Prices and sales for camcorders



Dynamic Demand

- Today a 55" 4K LCD TV is $400. In 2006, you could buy a 32" 720P TV for $> \$10,000$.
- In December 2011 TV prices fell 17% on an annual basis and other A/V equipment fell 11%, and computer equipment fell 14%.
- From August 2005 to August 2015 prices declined by 87.2%.
- We might also find that over time consumers buy better cameras or larger TV's
- The BLS tries to do *chaining* and *quality adjustments* but in high-tech products this can be very difficult.
- This has a potentially large impact on price indices (a small bias in the CPI can be billions of dollars in SSA/Medicare payments).
- Depending on the time period we might find that demand slopes upwards (lower prices lead to lower sales)

Dynamic Demand

Each consumer type is subscripted by i, and chooses a product j in period t to maximize utility:

$$\begin{aligned} u_{ijt} &= \underbrace{\alpha_i^x x_{jt} + \xi_{jt}}_{f_{jt}} - \alpha_i^p p_{jt} + \varepsilon_{ijt} \\ u_{i0t} &= f_{i0t} + \varepsilon_{i0t} \end{aligned}$$

In the static model we have $f_{i0t} = 0$ for every i, t.

- Where does the bias arise from?
- Correlation between f_{i0t} and prices.

We can think about dynamic models as having time varying utility for the outside option f_{i0t}

- What kind of good do I own now?
- What do I anticipate in the future? (prices and quality)

Outside Good

Goal is to endogenize the utility of the outside option f_{i0t}:

- Previous choice(s): $f_{ij,t-1}$ (depreciated?)
- Consumer's best option from tomorrow's market

Ad-Hoc approach:

- Just proxy with a time trend (Lou Prentice Ying 2012), (Eizenberg 2011) etc. That is $f_{i0t} = \gamma_{0i} + \gamma_{1i} t + \gamma_{2i} t^2 + \dots$
- We can get the elasticity correct.
- Not structural! Not helpful if we want to do counterfactuals! Can't get the elasticities under different conditions.

Assumptions

For our dynamic model to make sense we may want to place some restrictions on f_{i0t}:

- Rational Expectations: $E[f_{i0,t+1}|\Omega_t] = f_{i0,t+1}$.
- Dynamic Consistency
- Law of motion for consumer types: $w_{i,t+1} = h(w_{i,t}, s_{ijt})$
- We can formally write down a dynamic programming problem that consumers solve:

$$\begin{aligned} V_i(f_{i0t}, \varepsilon_{it}, \Omega_t) &= \max\{f_{i0t} + \beta E_\Omega[E_\varepsilon V_i(f_{i0t}, \varepsilon_{it}, \Omega_{t+1})|\Omega_t], \\ &\quad \max_j f_{ijt} - \alpha_i p_{jt} + \beta E_\Omega[E_\varepsilon V_i(f_{ijt}, \varepsilon_{it}, \Omega_{t+1})|\Omega_t]\} \end{aligned}$$

Replacement Problem

This Bellman has defined a *Replacement Problem*.

- You own a single durable good with the option to *upgrade* each period.
- When you upgrade you throw away the old durable and get nothing in exchange.
- After a purchase j you receive flow utility $f_{i0t+1} = f_{ijt}$ each period if you don't make a new purchase.
- We could add in depreciation if we wanted to.
- Other studies of durables have focused on *Secondary Markets* or *Perfect Rental Markets*.
- Resale: Important for cars, not relevant for high-tech.

Inclusive Value

Helpful to write: $EV_i(\Omega_t) = \int V_i(\varepsilon_{it}, \Omega_t) f(\varepsilon)$ Rust's Trick

$$\begin{aligned} V_i(f_{i0t}, \varepsilon_{it}, \Omega_t) &= \max\{f_{i0t} + \beta E_\Omega[EV_i(f_{i0t}, \Omega_{t+1})|\Omega_t] + \varepsilon_{i0t}, \\ &\quad \max_j f_{ijt} - \alpha_i p_{jt} + \beta E_\Omega[EV_i(f_{ijt}, \Omega_{t+1})|\Omega_t] + \varepsilon_{ijt}\} \end{aligned}$$

We can write the ex-ante expected utility of purchasing in period t without having to condition on which good you purchase:

$$\begin{aligned} \delta_i(\Omega_t) &= E_\varepsilon[\max_j f_{ijt} - \alpha_i p_{jt} + \beta E_\Omega[EV_i(f_{ijt}, \Omega_{t+1})|\Omega_t] + \varepsilon_{ijt}] \\ &= \log\left(\sum_j \exp[f_{ijt} - \alpha_i p_{jt} + \beta E_\Omega[EV_i(f_{ijt}, \Omega_{t+1})|\Omega_t]]\right) \end{aligned}$$

Inclusive Value Sufficiency

$$EV_i(f_{i0},\Omega) = \log\left(\exp[f_{i0} + \beta E_{\Omega'}[EV_i(f_{i0},\Omega')|\Omega]] + \exp(\delta_i(\Omega))\right) + \eta$$

where $\eta = 0.577215665$ (Euler's Constant).

The fact that the expected value function depends recursively on itself and $\delta_i(\Omega_t)$ (Inclusive Value) leads to the following assumption.

Inclusive Value Sufficiency

If $\delta_i(\Omega) = \delta_i(\tilde{\Omega})$ then $g(\delta_i(\Omega')|\Omega) = g(\delta_i(\tilde{\Omega}')|\tilde{\Omega})$ for all $\Omega, \tilde{\Omega}$.

- The idea is that δ tells me everything about the future evolution of the states
- More restrictive than it looks. δ is low because quality is low? or because prices are high? Is this the result of a dynamic pricing equilibrium? (No!)

Inclusive Value Sufficiency

Under IVS the problem reduces to

$$
\begin{aligned}
EV_i(f_{i0}, \delta_i) &= \log\left[\exp(f_{i0} + \beta E_{\Omega'}[EV_i(f_{i0}, \delta_i')|\delta_i]) + \exp(\delta_i)\right] \\
\delta_i &= \log\left(\sum_j \exp[f_{ijt} - \alpha_i p_{jt} + \beta E_{\delta'}[EV_i(f_{ijt}, \delta_i')|\delta_i]\right)
\end{aligned}
$$

The idea is that the inclusive value δ_{it} IS the state space, along with his current holding of the durable f_{i0t}.

Rational Expectations

We still have the expectation to deal with:

$$E_{\delta'}[EV_i(f_{ijt}, \delta_i')|\delta_i]$$

We need to take a stand on $g_i(\delta_i'|\delta_i)$ the anticipated law of motion for δ_i. G&R assume it follows an $AR(1)$ process.

$$\delta_{it+1} = \gamma_0 + \gamma_1 \delta_{it} + \nu_{it} \text{ with } \nu_{it} \sim N(0, \sigma_\nu^2)$$

If we see δ_{it} we could just run the $AR(1)$ regression to get consumer belief's $\hat{\gamma}$

Rational Expectations-Interpolation

I still haven't told you how to compute

$$E_{\delta'}[EV_i(f_{ijt}, \delta_i')|\delta_i, \gamma] = \int EV_i(f_{ijt}, \delta_i')g(\delta'|\delta, \gamma)$$

1. We need to integrate $EV(f_{ijt}, \delta_i)$ (a function) over a normal density.
2. But we don't observe $EV(f_{ijt}, \delta_i)$ everywhere, only on the grid points of our state space.
3. We can fit a linear function, cubic spline, etc. over δ_i to EV_i at each value of f_{ijt} on our grid.
4. We need to interpolate $\widehat{EV}_i(\delta_i^s)$ (Linear, Cubic Spline, etc.)
5. We might as well interpolate the function at the *Gauss-Hermite* quadrature nodes and weights, recentered at $\gamma_0 + \gamma_1 \delta$ in order to reduce the number of places we interpolate $\widehat{EV_i}$.

Rational Expectations-Alternative

There is an alternative method that is likely to be less accurate

$$E_{\delta'}[EV_i(f_{ijt}, \delta_i')|\delta_i, \gamma] = \int EV_i(f_{ijt}, \delta_i') g(\delta'|\delta, \gamma)$$

1. We need to integrate $EV(f_{ijt}, \delta_i)$ (a function) over a normal density but we only see it at the grid points of our state space.
2. We could discretize $g(\delta'|\delta, \gamma)$ so that it is a valid markov transition probability matrix (TPM) evaluated only at the grid points.
3. Now computing the expectation is just matrix multiplication.

I am a bit nervous about whether two discrete approximations will get the continuous integral correct.

The Estimation Problem

We need to solve $\forall i, t$:

$$
\begin{aligned}
S_{jt} &= \sum_i w_i s_{ijt}(f_{i0t}, \delta_{it}) \\
f_{ijt} &= \overline{\alpha} x_{jt} + \xi_{jt} + \sum_l \sigma_l x_{jl} \nu_{il} \\
s_{ijt}(f_{i0t}, \delta_{it}) &= \frac{\exp[f_{ijt} - \alpha_i p_{jt} + \beta E_{\Omega'}[EV_i(f_{ijt}, \delta_i')|\delta_i]}{\exp[EV_i(f_{i0t}, \delta_{it})]} \\
EV_i(f_{i0}, \delta_i) &= \log\left[\exp(f_{i0} + \beta E_{\Omega'}[EV_i(f_{i0}, \delta_i')|\delta_i]) + \exp(\delta_i)\right] \\
\delta_i(EV_i) &= \log\left(\sum_j \exp[f_{ijt} - \alpha_i p_{jt} + \beta E_{\delta'}[EV_i(f_{ijt}, \delta_i')|\delta_i]\right) \\
E[\delta_{it+1}|\delta_{it}] &= \gamma_0 + \gamma_1 \delta_{it} \\
w_{i,t+1} &= h(w_{i,t}, s_{ijt})
\end{aligned}
$$

The Estimation Problem

1. Like BLP we guess the nonlinear parameters of the model θ
2. For a guess of the ξ_{jt}'s we can solve for EV_i by iteratively computing δ, and running the γ regression for each i and spline/interpolating to compute $E[EV_i]$. (Inner Loop)
3. G&R show how the contraction mapping of BLP can be modified to find a fixed point of the δ, ξ, γ relationship to find f_{ijt} (Middle Loop).
4. We need to make sure to update the $w_{i,t}$ via $h(\cdot)$. (This is a TPM that tells maps the transition probabilities of type i holding f_{i0t} to $f_{i0,t+1}$).
5. Once we've solved this whole system of equations, we use ξ to form moments just like BLP and do GMM. (Outer Loop)

G&R Parameters

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	.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) *

G&R Robustness

Table 2: Robustness

Parameter	State space includes number of products	Perfect foresight	Dynamic model with extra random coefficients	Linear price	Melnikov's model	Month dummies
	(1)	(2)	(3)	(4)	(5)	(6)
Mean coefficients (α)						
Constant	-.098 (.026) *	-.129 (.108)	-.103 (.037) *	-.170 (.149)	-6.61 (.815) *	-.114 (.024) *
Log price	-3.31 (1.04) *	-2.53 (.940) *	-3.01 (.717) *	-6.94 (.822) *	-.189 (.079) *	-3.06 (.678) *
Log size	-.007 (.001) *	-.006 (.001) *	-.015 (.007) *	.057 (.008) *	-.175 (.049) *	-.007 (.001) *
Log pixel	.010 (.003) *	.008 (.001) *	.009 (.002) *	.037 (.012) *	-.288 (.053) *	.010 (.002) *
Log zoom	.005 (.002) *	.004 (.002) *	.004 (.002)	-.117 (.012) *	.609 (.074) *	.005 (.002)*
Log LCD size	.004 (.002) *	.004 (.001) *	.004 (.002) *	.098 (.010) *	-.064 (.088)	.003 (.001) *
Media: DVD	.033 (.006) *	.025 (.004) *	.044 (.018) *	.211 (.053) *	.147 (.332)	.031 (.005) *
Media: tape	.013 (.005) *	.010 (.004) *	.024 (.016)	.200 (.051) *	-.632 (.318) *	.012 (.004) *
Media: HD	.036 (.009) *	.026 (.005) *	.047 (.019) *	.349 (.063) *	-.545 (.419)	.034 (.007) *
Lamp	.005 (.002) *	.003 (.001) *	.005 (.002) *	.077 (.011) *	-.200 (.058) *	.004 (.001) *
Night shot	.003 (.001) *	.004 (.001) *	.003 (.001) *	-.062 (.008) *	.427 (.058) *	.003 (.001) *
Photo capable	-.007 (.002) *	-.005 (.002) *	-.007 (.002) *	-.061 (.019) *	-.189 (.142)	-.007 (.008)
Standard deviation coefficients ($\Sigma^{1/2}$)						
Constant	.085 (.019) *	.130 (.098)	.081 (.025) *	.022 (.004) *		.087 (.013) *
Log price	.349 (.108) *	2.41e-9 (.919)	1.06e-7 (.522)	1.68 (.319) *		.287 (.078) *
Log size			-.011 (.007)			
Log pixel			1.58e-10 (.002)			

Standard errors in parentheses; statistical significance at 5% level indicated with *. All models include brand dummies, with Sony excluded.
There are 4436 observations, except in the yearly model, in which there are 505.

Perfect Foresight

- Contrary to the static model, price coefficient is negative (as one would expect).
- Coefficients on many product characteristics are intuitively appealing.
- Allowing for repeated purchases generates more "sensible" results.
- "Better results" from a dynamic model may be due to the fact that people wait to purchase because of the expectations of price declines and not directly because of high prices.
- Unlike the static model, in dynamic setup the explanation of waiting does not conflict with consumers buying relatively high-priced products.
- A variety of robustness measures show that the major simplifying assumptions about the dynamics in the model are broadly consistent with the data.

Perfect Foresight

G&R Report similar elasticities in the perfect foresight case. We make the following simplification

$$E_{\Omega'}[EV_i(f_{i0}, \delta_{i,t+1})|\delta_{i,t}])] = EV_i(f_{i0}, \delta_{i,t+1})$$

This saves us a lot of headaches:

- No more integration/interpolation
- We can solve the problem on the grid!
- No more belief regressions

Conlon (2014)

Conlon (2014) suggests an alternate formulation of the problem.

Assume briefly that there are no upgrades $f_{i0t} = 0$, as well as perfect foresight:

$$E_{\Omega'}[EV_i(f_{i0}, \delta_{i,t+1})|\delta_{i,t}])] \quad = \quad EV_i(f_{i0}, \delta_{i,t+1}) = v_{it}$$

$$v_{it} \equiv EV_i(f_{i0}, \delta_i) \quad = \quad \log\left[\exp(\beta v_{i,t+1}) + \exp(\delta_i)\right] + \eta$$

The Estimation Problem

We need to solve $\forall i, t$:

$$
\begin{aligned}
S_{jt} &= \sum_i w_i s_{ijt} \\
f_{ijt} &= \overline{\alpha} x_{jt} + \xi_{jt} + \sum_l \sigma_l x_{jl} \nu_{il} \\
s_{ijt} &= \exp[f_{ijt} - \alpha_i p_{jt} - v_{it}] \\
v_{it} &= \log\left[\exp(\beta v_{i,t+1}) + \exp(\delta_{it})\right] + \eta \\
\delta_{it} &= \log\left(\sum_j \exp[f_{ijt} - \alpha_i p_{jt}]\right) \\
w_{i,t+1} &= w_{i,t}\left(1 - \sum_j s_{ijt}\right)
\end{aligned}
$$

Note: that this nests static demand $\beta = 0$

Alternative Perspective on Beliefs

Recall our objective:

- Plug in an unbiased estimate for the “no-purchase” utility.
- Under perfect foresight this is just the inclusive value of tomorrow’s market $\delta_{i,t+1}$ appropriately discounted: $\sum_{k=1}^{T-k} \beta^{t+k} \delta_{i,t+k}$.
- Different ways to think about rational expectations
  - Expectational error of some or all of $\delta_{i,t+k}$'s.
  - Expectational error in today’s reservation utility.

Endogeneity and Instruments

- Dynamics mean we lean harder on the assumption of exogenous product characteristics
- In one period we can take characteristics as given, but in many periods this becomes less palatable (Do cameras exogenously improve over time?).
- Endogeneity: price is endogenous while other product characteristics are not, i.e. x_{jt}. (Size,Resolution, etc.)
- Price is chosen by the firms possibly after observing ξ_{jt} and, hence, is endogenous.
- Instruments: use variables that affect the price-cost margin, e.g. measures of how crowded a product is in characteristics space, which effects price-cost margin and the substitutability across products.
  1. all of the product characteristics in x;
  2. mean product characteristics for a given firm;
  3. mean product characteristics for all firms;
  4. the count of products offered by the firm and by all firms.
  5. changes in costs over time?

Hendel and Nevo (2006)

- When a supermarket cuts the price of laundry detergent for a week there is a huge increase in sales.
- This leads us to conclude consumers are extremely elastic with respect to price
- When a supermarket makes a permanent price cut to laundry detergent, there is little sales impact in the long run.
- Now consumers look highly inelastic with respect to price
- Often we use average prices which include high and low periods in regression studies – does this make sense?
- How can we resolve this puzzle?

Hendel and Nevo (2006)

- Hendel and Nevo suggest that consumers respond by temporary price reductions by stockpiling inventories.
- Consumers spend down their inventories during periods of high prices
- Consumers have variable storage costs and price sensitivities. Why?
- This has implications for inter temporal price discrimination and retail High-Low pricing strategies.

Data

- 9 Supermarkets in a large midwest city (Dominick's in Chicago)
- Store-level: for each brand 13 (j) size x: 32-256oz in each store, each week (t)
  1. Price p_{jxt}
  2. Quantity q_{jxt}
  3. Promotions a_{jxt} (binary for feature/display)
- Consumers are of type h with utility: $u(c_{ht} + \nu_{ht}; \theta_h)$
- Current consumption is $c_{ht} = \sum_j c_{jht}$ not brand specific!
- There is a shock affected marginal utility of consumption ν_{ht}.
- Decision: $d_{hjxt} = 1$ is a purchase of h of brand j and size x at t. (includes outside option $= 0$).

Table 3: Sales

	Quantity Discount (%)	Quantity Sold on Sale (%)	Weeks on Sale (%)	Average Sale Discount (%)	Quantity Share (%)
Liquid					
32 oz.	—	2.6	2.0	11.0	1.6
64 oz.	18.1	27.6	11.5	15.7	30.9
96 oz.	22.5	16.3	7.6	14.4	7.8
128 oz.	22.8	45.6	16.6	18.1	54.7
256 oz.	29.0	20.0	9.3	11.8	1.6
Powder					
32 oz.	—	16.0	7.7	14.5	10.1
64 oz.	10.0	30.5	16.6	12.9	20.3
96 oz.	14.9	17.1	11.5	11.7	14.4
128 oz.	30.0	36.1	20.8	15.1	23.2
256 oz.	48.7	12.9	10.8	10.3	17.3

Dynamic Discrete Choice

$$
\begin{aligned}
V(s_t) &= \max_{c_h(s_t), d_{jxt}(s_t)} \sum_t \beta^{t-1} E[u(c_{ht} + \nu_{ht}; \theta_h) - C_h(i_{h,t+1}; \theta_h) \\
&+ \sum_j d_{hjxt}(\alpha_h^p p_{jxt} + \xi_{hjx} + \alpha_h^a a_{jxt} + \epsilon_{hjxt}) | s_t] \\
i_{h,t+1} &= i_{ht} + x_{ht} - c_{ht} \\
\sum_{j,x} d_{hjxt} &= 1
\end{aligned}
$$

- Abuse of notation: x_{ht} is size of the choice
- $C_h(i; \theta_h)$ is cost of storage
- s_t contains current inventory i_t, current prices, and consumption shock ν_t as well as ϵ_{ht}.
- ξ_{jhxt} captures expected future differences in utility of x units of j at time of purchase.
    1. as long as discounting is low
    2. brand-specific differences in utility (but not consumption) enter linearly.

Model Assumptions

Assumption 1

ν_t is independently distributed over time and across consumers.
No serial correlation!

Assumption 2

Prices p_{jxt} and advertising a_{jxt} follow an exogenous first-order Markov process.
Hard to justify this with a model of profit maximizing supply!

Assumption 3

ϵ_{jxt} is i.i.d. extreme value type 1.

Likelihood

Conditional on Household we can write the probability of a sequence of purchase decisions:

$$P(d_1, \ldots, d_t | p_1, \ldots, p_T) = \int \prod_t P(d_t | p_t,$$
$$i_t(d_{t-1}, \ldots, d_1, \nu_{t-1}, \ldots, \nu_1, i_1)) dF(\nu_1, \ldots, \nu_T) dF(i_1)$$

- Beginning of period inventory depends on previous decisions, previous shocks, and initial inventory.
- p_t now includes all observed state variables not just prices

Choice Problem

$$
\begin{aligned}
Pr(d_{jx}|p_t, i_t, \nu_t) &= \frac{\exp[\alpha p_{jxt} + \xi_{jx} + \beta a_{jxt} + M(s_t, j, x)]}{\sum_{k,y} \exp[\alpha p_{kyt} + \xi_{jy} + \beta a_{kyt} + M(s_t, k, y)]} \\
M(s_t, j, x) &= \max_c [u(c + \nu_t) - C(i_{t+1}) + \beta E[V(s_{t+1}|d_{jx}, c, s_t]]
\end{aligned}
$$

- State space has very high dimension. (Lots of brand-size combos at different prices)
- Keeping track of all brands/prices would be very costly

3-step Procedure

To reduce complexity, Hendel and Nevo propose a 3-step estimator

- Maximize likelihood of observed brand choice conditional on size in order to recover the (α, ξ) parameters.
- This avoids solving MDP but instead is just static discrete choice problem (efficiency loss!)
- Second step: compute inclusive values for each size and transition probability matrix.
- Now solve a quantity choice only nested fixed point problem. The key is that there is only one "index price" per size.
- The reason this is feasible is our old friend, the conditional independence assumption (of what?)

Step 1: Brand Choice

$$\begin{aligned} Pr(d_{jx}|x_t, p_t, i_t, \nu_t) &= \frac{\exp[\alpha^p p_{jxt} + \xi_{jx} + \alpha^a a_{jxt}]}{\sum_{k,y} \exp[\alpha^p p_{kyt} + \xi_{jy} + \alpha^a a_{kyt}]} \\ &= Pr(d_{jx}|x_t, p_t) \end{aligned}$$

- The trick is that $M(s_t, j, x)$ is the same for all products of the same size x.
- This means the dynamics drop out of the brand-choice equation conditional on x_t.
- We can recover (α, ξ) from static demand estimation!

Step 2: Inclusive Values

$$\omega_{xt} = \log\left(\sum_k \exp(\alpha^p p_{kxt} + \xi_{xt} + \alpha^a a_{kxt})\right)$$

Assumption 4: IVS

$$F(\omega_t | s_{t-1}) = F(\omega_t | \omega_{t-1})$$

- Compute ex-ante expected utility of purchasing size x in period t
- Does not depend on which j is purchased.
- IVS means we can keep track of a lot less information!
- Same as G&R two price vectors with same inclusive values must have same transition probabilities.
- Do individual prices still matter? (Test)

Step 3: Dynamic Choice of Size

$$V(i, \omega_t, \epsilon_t, \nu_t) = \max_{c,x}[u(c + \nu_t) - C(i_{t+1}) + \omega_{xt} + \epsilon_{xt} + \beta E[V(i_{t+1}, \omega_{t+1}, \epsilon_{t+1}, \nu_{t+1})|i_t, \omega_t, \epsilon_t, \nu_t, c, x]]$$

- Compute ex-ante expected utility of purchasing size x in period t
- Does not depend on which j is purchased.
- IVS means we can keep track of a lot less information!
- Same as G&R two price vectors with same inclusive values must have same transition probabilities.
- Do individual prices still matter? (Test)

Key Proposition

How do we know that the simplified problem has the same solution as the original dynamic problem?

$$P(x_t|i_t, p_t, \nu_t) = P(x_t|i_t, \omega(p_t), \nu_t)$$

- Before we got to see the entire state s_t
- Now we only see the expected utility of x_t aka ω_{xt}
- The proof relies on Assumption 3(IID Logit errors) and Assumption 4 (IVS).

Computational Details

Iterate policy evaluation and policy improvement

1. Approximate the value function by a polynomial function of s_t. (Logarithmic)
2. Guess an optimal policy and minimized (LSQ) deviation between the value function and expected future value
3. Update the policy function for every state
4. Update expectation with coefficients and expected value of state variables
5. Repeat until value function coefficients converge

This is a Smooth Approximation approach

Table 4: Brand Choice Estimates

	(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									✓	✓

Table 5: Belief Process Estimates

	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)

Table 5: Dynamic Problem Estimates

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

Table 8: Elasticities Compared to Static Model

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.