Switching Costs and Network Effects

C.Conlon

Grad IO

Fall 2016

State Dependence

Think about a static model like BLP

$$u_{ijt} = \beta_i x_{jt} - \alpha_i p_{jt} + \xi_{jt} + \epsilon_{ijt}$$

- Suppose I have panel data on consumer i's purchases and I observe that the consumer chooses different brands over time
- Why do you switch brands?
  1. New $\epsilon \rightarrow$ not helpful!
  2. Price responses $\rightarrow$ may wrongly attribute all effects to price.
  3. ξ_{jt} not correlated across individuals but may include things like advertising, etc.
- Challenge is explaining both persistence and switching behavior.

Why Do We Care?

- Switching costs appear to be a real friction in the economy.
- Consumers are often highly persistent in product choices.
  - Because they really like the product?
  - Because they are unaware of alternatives?
  - Because they are lazy?
- Extremely important in the market for health insurance. Consumers in ACA (Obamacare) exchanges would have saved $610/yr on average if they switched to a lower cost plan in the same tier.
  - Real costs associated with switching: checking to see if my doctor takes the other insurer, calculating expected expenditures, etc.
- Can we reduce or exploit frictions with laws? defaults? etc.

Why Do We Care?

- Switching costs are another way to escape the Bertrand trap for firms which sell relatively undifferentiated products.
- Old idea going back to Klemperer (1995), Farrell and Klemperer (2007). Do switching costs make markets more or less competitive?
- Two incentives:
  - Investment: Sign up a bunch of consumers today and they will be "sticky" to you in the future $\rightarrow$ lower prices
  - Harvesting: You have additional market power over your "sticky" customers $\rightarrow$ higher prices
- Most people believe that harvesting dominates, and switching costs lead to higher prices. (But not always...)

Cabral (JMR 2008)

Consider dynamic optimization problem faced by firm i with a vector of prices $\mathbf{p}$ and state variables (shares) $\mathbf{x}$ and switching costs s:

$$V_i(\mathbf{x},\mathbf{p},s) = (p_i - c_i)\cdot q_i(\mathbf{x},\mathbf{p},s) + \beta \tilde{V}_i(\mathbf{x},\mathbf{p},s)$$

with FOC

$$q_i(\mathbf{x},\mathbf{p},s) + (p_i - c_i)\cdot \underbrace{\frac{\partial q_i(\mathbf{x},\mathbf{p},s)}{\partial p_i}}_{q_i'} + \beta \underbrace{\frac{\partial \tilde{V}_i(\mathbf{x},\mathbf{p},s)}{\partial p_i}}_{\tilde{V}_i' \frac{\partial q_i}{\partial p_i}}$$

Define $\tilde{V}_i' \equiv \frac{\partial \tilde{V}_i}{\partial q_i}$ (note w.r.t. q_i not p_i). So that:

$$p_i - c_i = \underbrace{\frac{q_i}{-q_i'}}_{\text{Harvesting}} - \underbrace{\beta \tilde{V}_i'}_{\text{Investment}}$$

Cabral (JMR 2008)

$$p_i - c_i = \underbrace{\frac{q_i}{-q_i'}}_{\text{Harvesting}} - \underbrace{\beta \tilde{V}_i'}_{\text{Investment}}$$

- Second term (dynamic benefit of increasing q_i today) is "investing" in marketshare and leads to lower PCM.
- First term is additional market power from switching costs and leads to higher PCM.
- Take derivatives w.r.t. s.
  - It is clear that $|q_i'|$ is decreasing in s. Higher switching costs increase static market power.
  - q_i is ambiguous across firms. (So net effect is ambiguous across i).
  - V_i' should be zero if $s = 0$. And V_i' is increasing in s. (Always positive).
- Harvesting can be $\pm$, Investment always $-$.

How do we model these?

$$u_{ijt} = \beta_i x_{jt} - \alpha_i p_{jt} + \xi_{jt} + \gamma_i \cdot I[y_{i,t-1} = j] = \epsilon_{ijt}$$

- We can include lagged choice in utility of the agent. (First order Markov)
- Could include two lagged choices if we wanted to.
- Consumers are not forward looking. Why?
- Has some problems: endogeneity, correlation in ϵ_{ijt} over time, etc.
- Fundamental question: How do we identify separately from persistent brand preference?
- Dube, Histch, Rossi approach: Throw a ton of heterogeneity at the problem.

Mixture of Normals

Let $\theta_i = [\alpha_i, \beta_i, \gamma_i]$.

- For each individual draw a class k from a multinomial distribution π.
- Now draw $\theta_i \sim MVN(\mu_k, \Sigma_k)$.
- Idea is that $P(\theta_i|\pi, \mu, \Sigma) = \sum_k \pi_k \phi(\theta_i|\mu_k, \Sigma_k)$ is a mixture of normals.

Mixture of Normals

Let $\theta_i = [\alpha_i, \beta_i, \gamma_i]$.

- For each individual draw a class k from a multinomial distribution π.
- Now draw $\theta_i \sim MVN(\mu_k, \Sigma_k)$.
- Idea is that $P(\theta_i|\pi, \mu, \Sigma) = \sum_k \pi_k \phi(\theta_i|\mu_k, \Sigma_k)$ is a mixture of normals.
- These models are highly flexible (around 4-5 normals tends to well approximate most distributions).
- But hard to estimate! (Problem is highly non-convex, EM algorithm is slow).
- In order to do MCMC estimation we have to assume some hyper-parameters b so that we can put a prior on π as well as μ_k, Σ_k.

Switching Costs in Orange Juice

TABLE 1 **Data Description**

Product	Average Price ($)	Trips (%)
Margarine		
Promise	1.69	14.3
Parkay	1.63	5.4
Shedd's	1.07	13.8
I Can't Believe It's Not Butter!	1.55	25.6
No purchase		40.8
No. of households	429	
No. of trips per household	16.7	
No. of purchases per household	9.9	

Product	Average Price ($)	Trips (%)
Refrigerated orange juice		
64 oz Minute Maid	2.21	11.1
Premium 64 oz Minute Maid	2.62	7.0
96 oz Minute Maid	3.41	14.7
64 oz Tropicana	2.26	6.7
Premium 64 oz Tropicana	2.73	28.8
Premium 96 oz Tropicana	4.27	8.0
No purchase		23.8
No. of households	355	
No. of trips per household	12.3	
No. of purchases per household	9.4	

Switching Costs in Orange Juice

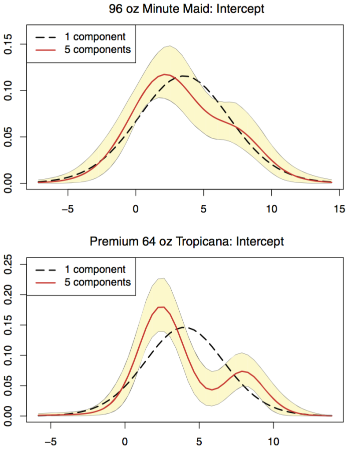
TABLE 2 Repurchase Rates

Product	Purchase Frequency	Repurchase Frequency	Repurchase Frequency after Discount
		Margarine	
Promise	.24	.83	.85
Parkay	.09	.90	.86
Shedd's	.23	.81	.80
ICBINB	.43	.88	.88
		Refrigerated orange juice	
Minute Maid	.43	.78	.74
Tropicana	.57	.86	.83

Switching Costs in Orange Juice

FIGURE 3

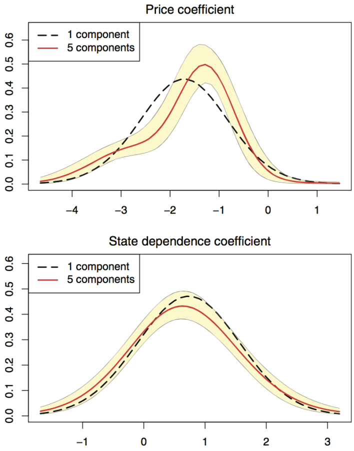
DISTRIBUTION OF BRAND INTERCEPTS: REFRIGERATED ORANGE JUICE



The graphs display the pointwise posterior mean and 90% credibility region of the marginal density of refrigerated orange juice brand intercepts (α_j^h). The results are based on a five-component mixture-of-normals heterogeneity specification. For comparison purposes, we also show the results from a one-component heterogeneity specification.

Switching Costs in Orange Juice

DISTRIBUTION OF PRICE AND STATE DEPENDENCE COEFFICIENTS:
REFRIGERATED ORANGE JUICE



The graphs display the pointwise posterior mean and 90% credibility region of the marginal density of the refrigerated orange juice price coefficient (η^h) and state dependence coefficient (γ^h). The results are based on a five-component mixture-of-normals heterogeneity specification. For comparison purposes, we also show the results from a one-component heterogeneity specification.

they also try 10 components...

Identification

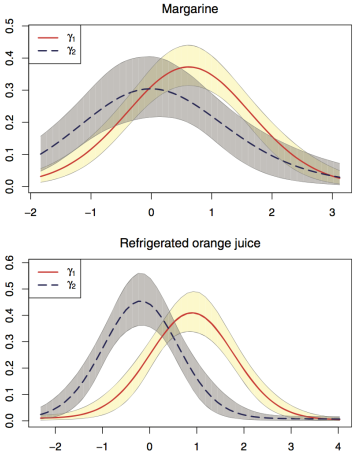
- Lots of price changes in the category. Imagine two brands (P, C) and each one can set two prices $\{H, L\}$.
- We observe the sequence $D_1(H, H) = C, D_2(H, L) = C, D_3(H, H) = C, D_4(L, H) = P$.
- If we see that $D_5(H, H/L) = P$ then we find evidence of state dependence.
- Likewise we can see you switch, become sticky, and switch back later.

Identification/Robustness

- The authors re-arrange the order of purchases within an individual and re-estimate.
- If this was persistent heterogeneity they should still spuriously find a large γ
- They do not!

Switching Costs in Orange Juice

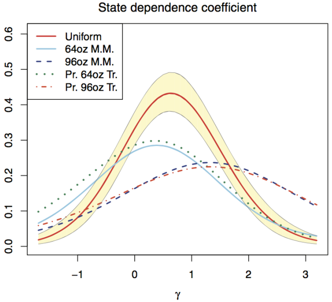
TESTING FOR AUTOCORRELATION



The graphs display the pointwise posterior mean and 90% credibility region of the marginal density of the coefficients γ_1 and γ_2 in model (12). γ_1 is the main state dependence coefficient, and γ_2 represents the effect of the interaction between the purchase state and the presence of a price discount when the product was last purchased. We expect that $\gamma_2 < 0$ under autocorrelated taste shocks. The results are based on a five-component mixture-of-normals heterogeneity specification.

Switching Costs in Orange Juice

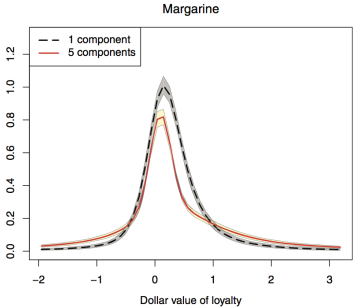
DISTRIBUTION OF BRAND-SPECIFIC STATE DEPENDENCE COEFFICIENTS: REFRIGERATED ORANGE JUICE



The graph displays the pointwise posterior mean and 90% credibility region of the marginal density of the state dependence coefficient (γ^h), based on a five-component mixture-of-normals heterogeneity specification. We show the densities both for a model specification with a uniform (across-brands) state dependence coefficient and for a specification allowing for brand-specific state dependence coefficients (we show results for the four orange juice brands with the largest market shares).

Switching Costs in Orange Juice

DISTRIBUTION OF THE DOLLAR VALUE OF LOYALTY MARGARINE



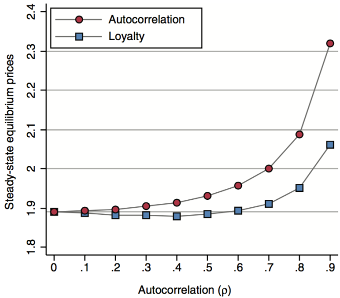
The graph displays the pointwise posterior mean and 90% credibility region of the marginal density of the dollar value of loyalty, defined as $-\gamma^h/\eta^h$. The results are based on a five-component mixture-of-normals heterogeneity specification. For comparison purposes, we also show the results from a one-component heterogeneity specification.

Why Does this matter

- Solve a dynamic programming problem like in Cabral (2008).
- If we have just auto-correlation and no switching costs, there is NO harvesting incentive.
- If we have switching costs than there is.
- Very small switching costs can make markets MORE competitive.

Switching Costs in Orange Juice

EQUILIBRIUM PRICES UNDER STATE DEPENDENCE AND AUTOCORRELATION



The graph displays the (symmetric) steady-state equilibrium prices from a model with autocorrelated random utility terms, and contrasts these “true” prices to the price predictions if the inertia in the brand choice data were attributed to structural state dependence in the form of loyalty.

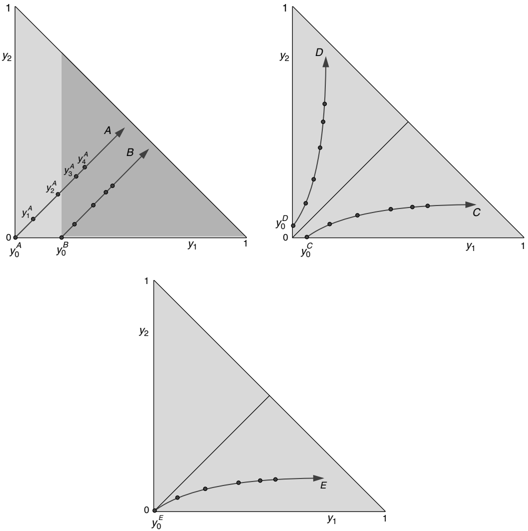
What are Network Effects?

- An important aspect of many digital markets today is *network effects*.
- Main idea is that you value the good more if other people use it.
  - Social Networks: Facebook, Instagram, Twitter, Tindr, etc.
  - Statistical Packages: Stata, R, Matlab, etc.
  - P2P Platforms: Ebay, Etsy, Alibaba, Uber.
  - Software Platforms: iOS, Android, Windows.
  - Game Consoles: PS4, XBox One, etc.
- This creates a lock in effect.
  - You may have an incentive to underprice initially to drive adoption.
  - There may be benefits to being early to market.
  - Markets can tip one way or another.
- Two-sided markets are another important issue (Developers, Developers, Developers!)

What are Network Effects?

- Consumers make adoption decision that is durable (or irreversible) and depends on two things:
  - The share of users on the same platform ρ_{jt}
  - Beliefs about the future of $E[\rho_{j,t}]$
- Because beliefs are important, multiple equilibria can arise
- How do we measure the size/impact of indirect network effects?
- Constructing a counterfactual equilibria in a world without network-effects is hard to do in practice.

What are Network Effects?



Dube, Hitsch, Chintagunta: Tipping

- Start with two firms and $M = 1$ mass of consumers
- Installed base $y_t = [y_{1t}, y_{2t}] \in [0, 1]$ is the state space.
- Assume that demand shock $\xi_{jt} \sim \phi(\xi)$ is private information to the firm (similar to Seim's paper on video stores).
- Timing of the game:
  1. Firms learn ξ_{jt} and set p_{jt}
  2. Consumers adopt $\{1, 2\}$ or delay purchase $= 0$
  3. Software firms supply a given number of titles n_{jt}
  4. Sales are realized and firms receive profits. Consumers receive utility from n_{jt} and in adoption period from platform itself.
- Information structure guarantees a unique best response (conjecture) and a pure-strategy equilibria.
- Hence prices p_{jt} contain a lot of information.

- Titles depend on next period state variable: $n_{jt} = h_j(y_{j,t+1})$. Why?

Consumers

Need two things:

- Current prices and installed base (p_t, y_t)
- Beliefs about the future $y_{t+1} = f^e(y_t, \xi_t)$ and conjecture about firm policy $p_{jt} = \sigma_j^e(y_t, \xi_t)$.

Utilities

- Flow from software: $u_j(y_{j,t+1}) = \gamma n_{jt} = \gamma h_j(y_{j,t+1})$
- In PDV: $\omega_j(y_{t+1}) = E[\sum_{k=0}^{\infty} \beta^k u_j(y_{j,t+1+k}) | y_{t+1}]$
- This PDV trick is common (and helpful) and solves the recursion:

$$\omega_j(y_{t+1}) = u_j(y_{j,t+1}) + \beta \int \omega_j(f^e(y_t, \xi_t))\phi(\xi)\partial\xi$$

Consumers

Choose j to maximize choice specific value function (indirect utility) logit error :

$$
\begin{aligned}
v_j(y_t, \xi_t, p_t) &= \delta_j + \omega_j(f^e(y_t, \xi_t)) - \alpha p_{jt} + \xi_{jt} \\
v_0(y_t, \xi_t) &= \beta \int \max\{v_0(y_{t+1}, \xi) + \varepsilon_0, \\
&\quad \max_j [v_j(y_{t+1}, \xi_t, \sigma^e(y_{t+1}, \xi)) + \varepsilon_j]\} \cdot \phi(\xi)\phi_\varepsilon(\varepsilon)
\end{aligned}
$$

This gives us logit shares $s_j(y_t, \xi_t, p_t)$ and a law of motion for y_t:

$$
y_{j,t+1} = y_{jt} + (1 - \sum_{k=1}^{J} y_{kt}) s_j(y_t, \xi_t, p_t) = f_j(y_t, \xi_t, p_t)
$$

Firms

- Constant marginal cost c_j and royalty rate r_j per unit of software $q_j(y_{t+1})$.
- Get $q_j(y_t)$ directly from the data.
- only integrate over your opponent's ξ_{-j}

$$\pi_j(y,\xi,p_j) = (p_j - c_j)\cdot(1-\sum_k^J y_{kt})\cdot\int s_j(y,\xi_j,\xi_{-j},p_j,\sigma_{-j}(y,\xi_{-j}))\phi_j(\xi_{-j})$$
$$+r_j\int q_j(f_j(y,\xi_j,\xi_{-j},p_j,\sigma_{-j}(y,\xi_{-j})))\phi_j(\xi_{-j})$$

Solve Bellman:

$$\begin{aligned} V_j(y,\xi_j) &= \sup_{p_j\geq 0}[\pi_j(y,\xi,p_j)+ \\ &\quad \beta_f\int V_j(f_j(y,\xi_j,\xi_{-j},p_j,\sigma_{-j}(y,\xi_{-j})))\phi(\xi_{-j})\phi(\xi_j')] \end{aligned}$$

Equilibrium

Define an MPE such that:

1. Choice specific value functions v_j and v_0 waiting value satisfy the Belmman Equation.
2. Firm's Value functions satisfy the Bellman equation
3. $p_j = \sigma_j(y, \xi_j)$ maximizes the RHS of the Bellman for each j in firm problem. (Tricky since econometrician doesn't see ξ directly).
4. Consumers have rational expectations $\sigma_j^e = \sigma_j$ and $f^e(y,\ \xi) = f(y, \xi, \sigma(y, \xi))$
5. Everyone acts rationally given expectations about the future, and those expecatations are consistent with what actually happens.

Data

- 32/64-bit console market , no backwards-compatibility, first to use CDROM
- 3DO had $\$700 - 1000$ console prices and failed to launch
- Sony Playstation was big winner: $9 royalty, low production cost.
- Sega Saturn was a failure. They exit console market completely afterwards
- N64 had lower console price but higher royalty $18. (and cartridge based)
- By Christmas of 1996 Nintendo had 8 games compared to PS 200.
- No must-buy title on PS.

Data and Estimates

Table 1 Descriptive Statistics

	Console	Mean	Std. dev.	Min	Max
Sales	PlayStation	275,409	288,675	26,938	1,608,967
	Nintendo	192,488	201,669	1,795	1,005,166
Price	PlayStation	119.9	30.3	55.7	200.6
	Nintendo	117.6	33.9	50.3	199.9
Game titles	PlayStation	594.2	381.1	3	1,095
	Nintendo	151.2	109.9	1	281

Table 2 Second-Stage Parameter Estimates

	Model 3		Model 7	
	Estimate	Std. error	Estimate	Std. error
δ_{Sony}	−1.21	0.89	−1.119	0.971
δ_{N64}	−1.34	0.87	−1.119	1.093
α	−1.94	0.52	−1.923	0.460
Time (<60)	−0.04	0.01	−0.049	0.028
γ (n_{jt}/1,000)	0.09	0.04	0.090	0.040
ψ (std. dev. of ξ_{jt})	0.05	0.09	0.028	1.950

Notes. Model 7 uses PPIs and exchange rates as instruments in first stage. $\beta = 0.9$; number of simulations $= 60$.

Counterfactual

Suppose we got rid of network effects, how much lower would the concentration of the market be?

Results

Table 3 Predicted One-Firm Concentration Ratios

Model predictions: Symmetric case (parameter estimates for Sony)				
Scale factor for γ	0.25	0.50	0.75	1.00
C_1	0.501	0.503	0.508	0.845
Discount factor (β)	0.600	0.700	0.800	0.900
$C_1{}^a$	0.501	0.502	0.508	0.845
$C_1{}^b$	0.501	0.501	0.508	0.845
Model predictions: Estimated parameter values				
Scale factor for γ	0.250	0.500	0.750	1.000
C_1	0.600	0.593	0.562	0.843
Discount factor (β)	0.600	0.700	0.800	0.900
$C_1{}^a$	0.602	0.601	0.599	0.843
$C_1{}^b$	0.571	0.572	0.562	0.843

Notes. The results are based on 5,000 simulations, and the concentration ratios are reported for month $T = 48$. No standard has an initial advantage; $y_0 = (0, 0)$.

[a]All estimated model parameters were obtained for $\beta = 0.9$.

[b]Predictions where the model parameters were reestimated for each consumer discount factor, β.

Table 4 Predicted Degree of Tipping at Estimated Parameter Values ($\beta = 0.9$)

Scale factor for γ	0.250	0.500	0.750
C_1	0.243	0.249	0.280
Discount factor (β)	0.600	0.700	0.800
$\Delta C_1{}^a$	0.241	0.242	0.244
$\Delta C_1{}^b$	0.272	0.271	0.280

Notes. This table displays the increase in market concentration relative to a specific counterfactual model, where either the marginal utility of software, γ, is scaled or a different consumer discount factor is chosen. The results are based on 5,000 simulations, and the tipping measures are reported for month $T = 48$. No standard has an initial advantage; $y_0 = (0, 0)$.

[a]All estimated model parameters were obtained for $\beta = 0.9$.

[b]Predictions where the model parameters were reestimated for each consumer discount factor, β.

Results

Table 5 Profit Increase for Installed Base Advantage

Installed base adv. of Sony	Discount factor (β) 0.6	0.7	0.8	0.9
0.025	70	134	370	808
0.050	139	271	732	1,142
0.075	207	410	1,052	1,271
0.100	274	547	1,317	1,381
0.125	339	680	1,529	1,470
0.150	403	807	1,711	1,541
0.175	464	922	1,857	1,589
0.200	523	1,030	1,985	1,617

Notes. This table shows the increase in the expected present discounted value of Sony's profits, measured in millions of dollars, for a given initial installed base advantage. The results are based on 5,000 simulations, and the present discounted value of profits is calculated for a time horizon of 48 months after the competitor (Nintendo) enters the market.