# Switching Costs: Orange Juice

C.Conlon

Fall 2024

Grad IO

# Additional Reading

You may want to take at the handbook chapter of Farrel and Klemperer for a review of the (largely theoretical) literature.

# State Dependence

Think about a static model like BLP

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

- 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? $\beta_i$ are persistent.
  1. New $\varepsilon \rightarrow$ not helpful!
  2. Price responses $\rightarrow$ may wrongly attribute all effects to price.
  3. $\xi_{jt}$ not correlated across individuals but may include things like advertising, etc.
- Challenge is explaining both persistence and switching behavior.

# Terminology

Sometimes we call these models switching costs and other times state dependence

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

- The idea is purchases in period $t-1$ have a causal effect on utility in period $t$
- We can think of this as either increasing utility for $j$ if you previously purchased it or providing an additional cost if $y_{it} \neq y_{i,t-1}$.

## 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 $\longrightarrow$ lower prices
  - Harvesting: You have additional market power over your “sticky” customers $\longrightarrow$ higher prices
- Most people believe that harvesting dominates, and switching costs lead to higher prices. (But not always...)

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?

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

- 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. Models are often time inconsistent. Why?
- Has some problems: endogeneity, correlation in $\epsilon_{ijt}$ 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.

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

- For each individual draw a class $k$ from a multinomial distribution $\pi$.
- 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 $\pi$ as well as $\mu_k, \Sigma_k$.

## Mixture of Normals

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

- For each individual draw a class $k$ from a multinomial distribution $\pi$.
- 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 $\pi$ as well as $\mu_k, \Sigma_k$.

# 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

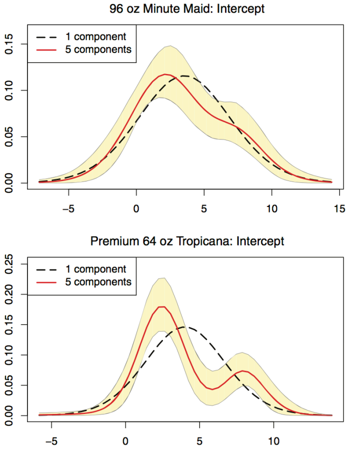
**TABLE 2 Repurchase Rates**

| 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

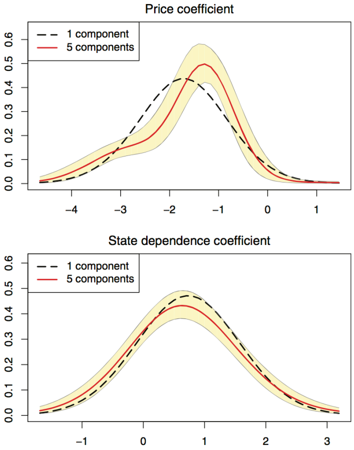
DISTRIBUTION OF BRAND INTERCEPTS: REFRIGERATED ORANGE JUICE




The graphs display the pointwise posterior mean and 90% credibility region of the marginal density of refrigerated orange juice brand intercepts ($\alpha_j^h$). 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



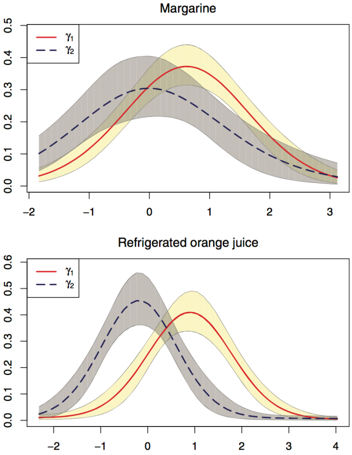

The graphs display the pointwise posterior mean and 90% credibility region of the marginal density of the refrigerated orange juice price coefficient ($\eta^h$) and state dependence coefficient ($\gamma^h$). The results are based on a five-component mixture-of-normals heterogeneity specification. For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

- 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.

- 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 $\gamma$
- They do not!

# Switching Costs in Orange Juice

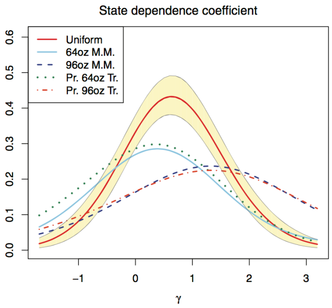
TESTING FOR AUTOCORRELATION




The graphs display the pointwise posterior mean and 90% credibility region of the marginal density of the coefficients $\gamma_1$ and $\gamma_2$ in model (12). $\gamma_1$ is the main state dependence coefficient, and $\gamma_2$ 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.

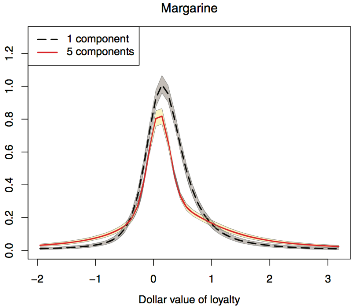
# Switching Costs in Orange Juice

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 ($\gamma^h$), 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).

DISTRIBUTION OF THE DOLLAR VALUE OF LOYALTY MARGARINE

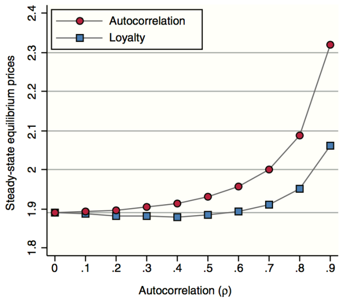



The graph displays the pointwise posterior mean and 90% credibility region of the marginal density of the dollar value of loyalty, defined as $-\gamma^h/\eta^h$. The results are based on a five-component mixture-of-normals heterogeneity specification. For comparison purposes, we also show the results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

- 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.