Switching Costs

C.Conlon

Fall 2019

Grad IO

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? β_i are persistent.
  1. New $\varepsilon \rightarrow$ 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.

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 $\rightarrow$ 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).

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. 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

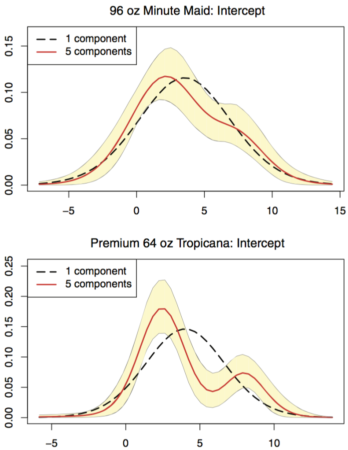
TABLE 2 Repurchase Rates

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

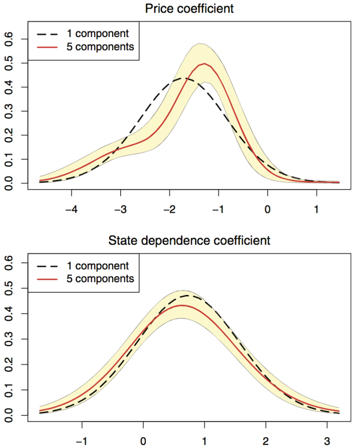
DISTRIBUTION OF BRAND INTERCEPTS: REFRIGERATED ORANGE JUICE



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.

Identification

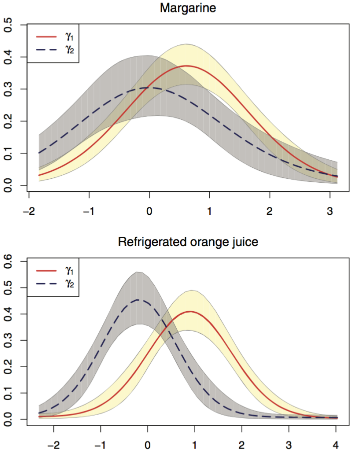
- Lots of price changes in the category. Imagine two brands (P, C) and each one can set two prices $\{H, L\}$.
- We observe the sequence $D_1(H, H) = C, D_2(H, L) = C, D_3(H, H) = C, D_4(L, H) = P$.
- If we see that $D_5(H, H/L) = P$ then we find evidence of state dependence.
- Likewise we can see you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

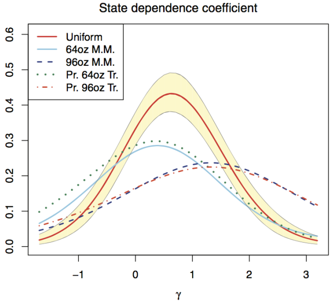
TESTING FOR AUTOCORRELATION



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

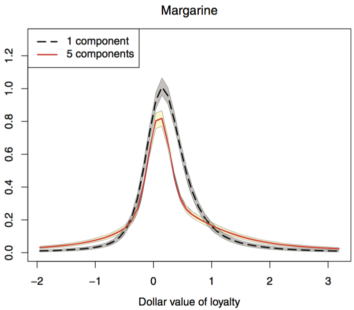
DISTRIBUTION OF BRAND-SPECIFIC STATE DEPENDENCE COEFFICIENTS: REFRIGERATED ORANGE JUICE



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



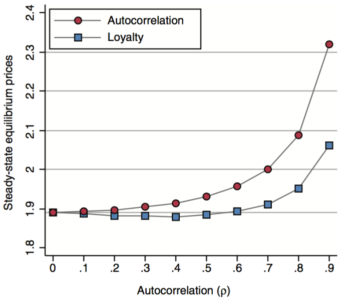
The graph displays the pointwise posterior mean and 90% credibility region of the marginal density of the dollar value of loyalty, defined as $-\gamma^h/\eta^h$. The results are based on a five-component mixture-of-normals heterogeneity specification. For comparison purposes, we also show the results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

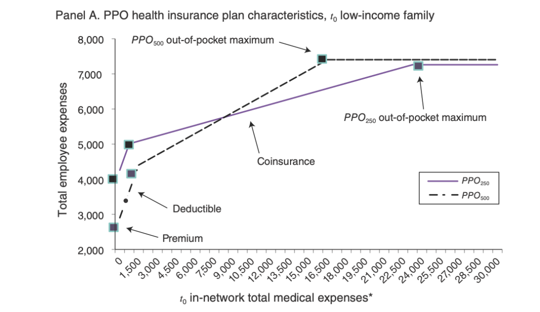
How do insurance contracts look?

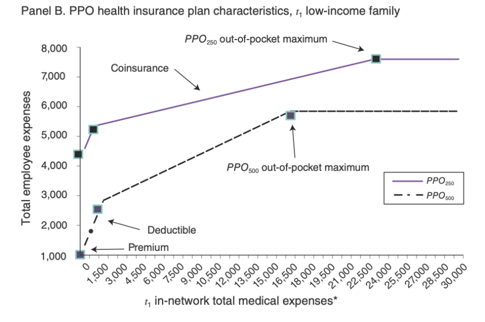
TABLE 1—DESCRIPTIVE STATISTICS

Sample demographics	All employees	PPO ever	Final sample
N–Employee only	11,253	5,667	2,023
N–All family members	20,963	10,713	4,544
Mean employee age (median)	40.1 (37)	40.0 (37)	42.3 (44)
Gender (male) percent	46.7	46.3	46.7
Income (percent)			
Tier 1 (< \$41K)	33.9	31.9	19.0
Tier 2 (\$41K–\$72K)	39.5	39.7	40.5
Tier 3 (\$72K–\$124K)	17.9	18.6	25.0
Tier 4 (\$124K–\$176K)	5.2	5.4	7.8
Tier 5 (> \$176K)	3.5	4.4	7.7
Family size (percent)			
1	58.0	56.1	41.3
2	16.9	18.8	22.3
3	11.0	11.0	14.1
4+	14.1	14.1	22.3
Staff grouping (percent)			
Manager (percent)	23.2	25.1	37.5
White-collar (percent)	47.9	47.5	41.3
Blue-collar (percent)	28.9	27.3	21.1
Additional demographics			
Quantitative manager (percent)	12.8	13.3	20.7
Job tenure mean years (median)	7.2 (4)	7.1 (3)	10.1 (6)
Zip code population mean (median)	42,925 (42,005)	43,319 (42,005)	41,040 (40,175)
Zip code income mean (median)	\$56,070 (\$55,659)	\$56,322 (\$55,659)	\$60,948 (\$57,393)
Zip code house value mean (median)	\$226,886 (\$204,500)	\$230,083 (\$209,400)	\$245,380 (\$213,300)

Notes: This table presents summary demographic statistics for the population we study. The first column describes demographics for the entire sample, whether or not they ever enroll in insurance with the firm. The second column summarizes these variables for the sample of individuals who ever enroll in a *PPO* option, the choices we focus on in the empirical analysis. The third column describes our final estimation sample, which includes those employees who (i) are enrolled in PPO_{-1} at t_{-1} and (ii) remain enrolled in any plan at the firm through at least t_1. Comparing the columns shows little selection on demographics into *PPO* options and some selection based on family size into the final estimation sample.

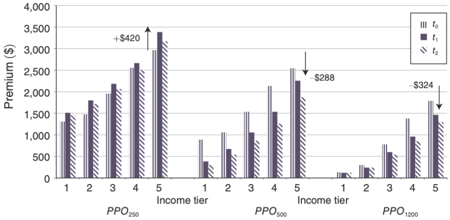
How do insurance contracts look?



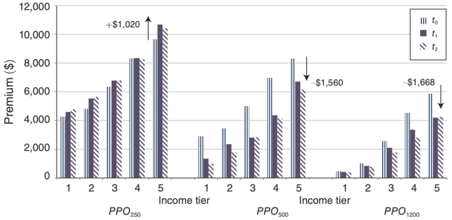


Evolution of Premiums

Panel A. PPO employee premiums, individual tier



Panel B. PPO employee premiums, family tier



Evidence of Switching Costs: New Employees and Dominated Plan

TABLE 2—NEW EMPLOYEE HEALTH PLAN CHOICES

New enrollee analysis	New enrollee t_{-1}	New enrollee t_0	New enrollee t_1
N, t_0	1,056	1,377	—
N, t_1	784	1,267	1,305
t_0 Choices			
PPO_{250}	259 (25%)	287 (21%)	—
PPO_{500}	205 (19%)	306 (23%)	—
PPO_{1200}	155 (15%)	236 (17%)	—
HMO_1	238 (23%)	278 (20%)	—
HMO_2	199 (18%)	270 (19%)	—
t_1 Choices			
PPO_{250}	182 (23%)	253 (20%)	142 (11%)
PPO_{500}	201 (26%)	324 (26%)	562 (43%)
PPO_{1200}	95 (12%)	194 (15%)	188 (14%)
HMO_1	171 (22%)	257 (20%)	262 (20%)
HMO_2	135 (17%)	239 (19%)	151 (12%)
Demographics			
Mean age	33	33	32
Median age	31	31	31
Female percent	56%	54%	53%
Manager percent	20%	18%	19%
FSA enroll percent	15%	12%	14%
Dental enroll percent	88%	86%	86%
Median (mean) expense t_1	844 (4,758)	899 (5,723)	—
Income tier 1	48%	50%	47%
Income tier 2	33%	31%	32%
Income tier 3	10%	10%	12%
Income tier 4	5%	4%	4%
Income tier 5	4%	5%	5%

Notes: This table describes the choice behavior of new employees at the firm over several consecutive years and presents our first model-free test of inertia. Each column describes one cohort of new employees at the firm, corresponding to a specific year of arrival. First, the chart describes the health insurance choices made by these cohorts in year t_0 (the year of the insurance plan menu change) and in the following year, t_1. The last part of the chart lists the demographics for each cohort of new arrivals at the time of their arrival. Given the very similar demographic profiles and large sample size for each cohort, if there is no inertia, the t_1 choices of employees who entered the firm at t_0 and t_{-1} should be very similar to the t_1 choices of employees who entered the firm at t_1. The table shows that, in fact, the active choices made by the t_1 cohort are quite different than those of the prior cohorts in the manner we would expect with high inertia: the t_1 choices of employees who enter at t_0 and t_{-1} reflect both t_1 prices and t_0 choices while the t_1 choices of new employees at t_1 reflect t_1 prices.

TABLE 3—DOMINATED PLAN CHOICE ANALYSIS

Dominated plan analysis	t_1 Dominated stay	t_1 Dominated switch	t_2 Dominated stay	t_2 Dominated switch
N	498	61	378	126
Minimum money lost[a]	$374	$453	$396	$306
PPO_{500}	—	44 (72%)	—	103 (81%)
PPO_{1200}	—	4 (7%)	—	6 (5%)
Any *HMO*	—	13 (21%)	—	17 (14%)
FSA t_1	25.4%	32.1%	27.2%	28.6%
FSA t_2	—	—	28.1%	30.9%
Dental switch t_1	4.3%	14.1%	3.5%	10.9%
Dental switch t_2	—	—	6.9%	17.2%
Age (mean)	44.9	38.3	46.2	41.4
Income tier (mean)[b]	1.6	1.4	1.6	1.7
Quant. manager	11%	8%	11%	11%
Single (percent)	40%	41%	40%	33%
Male (percent)	42%	46%	39%	55%
All plan analysis	PPO_{250} stay t_1	PPO_{250} switch t_1	All plans t_1 stay	All plans t_1 switch
Sample size	1,626	174	2,786	384
FSA t_1 enrollee	31%	41%	25%	39%
Dental switch	3.2%	13.1%	3.8%	14.5%
Age (mean)	48.3	40.6	44.0	39.1
Income tier (mean)[b]	2.5	2.2	2.3	2.1
Quant. manager	20%	17%	17%	14%
Single (percent)	50%	56%	53%	59%
Male (percent)	48%	42%	49%	40%

Notes: This top panel in this table profiles the choices and demographics of the employees enrolled in PPO_{250} at t_0 who (i) continue to enroll in a firm plan in t_1 and (ii) have PPO_{250} become dominated for them at t_1. The majority of these employees (498 out of 559 (89 percent)) remain in PPO_{250} even after it becomes dominated by PPO_{500} with 378 of 504 (25 percent) still remaining in this plan at t_2. People who do switch are more likely to exhibit a pattern of active choice behavior in general as evidenced by their higher FSA enrollments and level of dental plan switching. Apart from this, these populations are similar though switchers in this group are slightly younger. The bottom panel studies the profiles of those who switch at t_1 and those who don't for the two groups of (i) PPO_{250} enrollees at t_0 and (ii) the entire universe of PPO plan enrollees present in t_0 and t_1. This reveals a similar pattern of active decision making as switchers in these populations are also more likely to enroll in FSAs and switch dental plans.

Handel: Empirical Model

Use what Einav, Finkelstein, and Levin (2010) call a "realized" empirical utility model and assume that U_{kjt} has the following von-Neuman Morgenstern (v-NM) expected utility formulation

$$U_{kjt} = \int_0^\infty u_k\left(W_k, OOP, P_{kjt}, 1_{kj,t-1}\right) f_{kjt}(OOP)dOOP$$

$$u_k(x) = -\frac{1}{\gamma_k\left(\mathbf{X}_k^A\right)} e^{-\gamma_k\left(\mathbf{x}_k^A\right)_x}$$

- k is a family unit, j is an insurance plan, t is a year (t_0, t_1, t_2).
- $\gamma = \frac{u''(\cdot)}{u'(\cdot)}$ CARA risk-aversion (larger is more risk-averse).

Handel: Empirical Model

$$x = W_k - P_{kjt} - OOP + \eta\left(\mathbf{X}_{kt}^B, Y_k\right) 1_{kj,t-1} + \delta_k\left(Y_k\right) 1_{1200} + \alpha H_{k,t-1} 1_{250} + \epsilon_{kjt}\left(Y_k\right)$$

- W_k family wealth.
- P_{kjt} is the price for insurance plan j to family k.
- OOP is a draw from the distribution of $f(OOP)$ expenses: depends on the plan.
- $\eta\left(\mathbf{X}_{kt}^B, Y_k\right) 1_{kj,t-1}$ is the switching cost which depends on demographics $\mathbf{X}_{kt}^B$.
- $\delta_k(Y_k)$ is the family specific intercept for high-deductible plan (Y_k) is family dummy.
- $\alpha H_{k,t-1} 1_{250}$ is interaction between 90th percentile spenders and most generous plan.