Explain Consumer Heterogeneity in Structural State Dependence

Xiaoyuan Wang

School of Management and Economics, University of Electronic Science and Technology of China, Chengdu 611731, China

Email: xywang1987@gmail.com; wangxy@uestc.edu.cn

Explain Consumer Heterogeneity in Structural State Dependence

This paper investigates the structural state dependence effect in consumer packaged goods markets using scanner data sets. We use consumers' switch behaviour in a different product category and different years to provide possible inter-temporal preference variations that are independent of products and markets. Accounting for the household characteristics, we find that at least part of the variation in switch behaviour is stable: they explain a significant portion of consumers' state dependence in the market under investigation. Therefore, consumers' switch tendencies can be structural to their preference. The finding suggests that incorporating consumer switch behaviour from other choice domains can be a simple and effective way to learn consumers' heterogeneous structural state dependence.

Keywords: heterogeneity; brand switch; structural state dependence; scanner data

JEL: M31 D12 D90

Introduction

It is well documented in both marketing and economics literature that consumers' brand choices are subjected to the 'structural' state dependence effect (Dubé, Hitsch and Rossi 2010). Consumers are more likely to stay in their previous choices, not only because of product characteristics and market environments but also due to changes in consumers' intrinsic preference (e.g., due to habit formation). Unlike external factors, such changes are usually difficult to identify and thus capturing the heterogeneity in structural state dependence has been a challenge for both researchers and business practitioners.

The typical way of modelling heterogeneous consumer state dependence in the discrete choice framework (Rossi and Allenby 2005, Train 2009) is to allow the coefficients of the lagged choices to flexibly follow some pre-specified random distributions. To further rule out other sources of state dependence, such as market

frictions, consumer inertia studies usually carefully select markets in which confounding effects are minimal or can be controlled (Keane 1997, Seetharaman 2004, and Dubé, Hitsch and Rossi 2010). Those studies usually report positive and robust average structural state dependence effects, confirming that previous choices can change consumers' intrinsic preference.

The above empirical strategy is appropriate for a market with sufficient purchase history, flat learning curve, and little search cost. However, in reality, marketers would be more interested in evaluating consumer demand in a market with relatively inexperienced consumers and significant market frictions. In those markets, past choices usually cannot provide sufficient variation to distinguish heterogeneous state dependence from other unobserved temporal effects, leading to 'spurious state dependence' (Heckman 1981). Therefore, marketers have to rely on other sources of variations that may drive structural state dependence. In this paper, we test if variations in brand switches in a loosely related market may reveal structural state dependence. If a part of consumers' switch tendencies is truly intrinsic, we expect to see a strong interactive effect.

Two market categories, namely yoghurt (in Year 2001 or 2002) and carbonated beverage (in Year 2007), are selected from the IRI's scanner datasets (Bronnenberg, Kruger and Mela 2008). We introduce a simple measure, 'switch per choice' (SPC), to describe households' switch tendencies in the Yoghurt category. The switch information is then used in the models of (the same) households' brand choices in the carbonated beverage category.

Our estimation result shows that SPC in yoghurt markets can explain a significant portion of state dependence on major brand choices in the carbonated beverage markets. Such effects are robust when controlling for rich household level

demographic information. We conclude that SPC offers a practical way to capture some structural state dependence.

Data and Hypothesis Development

The IRI data consists of households' purchasing records in two mid-sized cities (Eau Claire, Wisconsin and Pittsfield, Massachusetts) of the United States. We model consumers' binary brand choices in the IRI carbonated beverage category in Year 2007; 4189 households have more than five shopping trips within this year.[1] In addition, we measure households' tendency to switch using the IRI yoghurt category from Year 2001 and 2002.

The IRI carbonated beverage market records about 43 vendor brands and 146 different sub-brands. We consider the choice decisions of the two main vendor brands: Coca-Cola and Pepsi, as well as all private label brands. Those three choices take approximately 80% of the total market share during year 2007. The IRI yoghurt market contains about 23 vendor brands and 89 different sub-brands. Brand switches are defined similarly at the vendor brand level in the yoghurt category: 10 major vendors[2] and private label brands take 96.5% of the total yoghurt market share.[3] Both markets witness frequent vendor level switches. For an average household in the yoghurt market or the carbonated beverage market, about 50% of the shopping trips involve a vendor brand switch. The summary statistics of vendor switches are provided in Table 1.

---

[1] In the IRI scanner data sets, choices and switches are defined at the household level. For expositional ease, we use 'household' and 'consumer' interchangeably.

[2] They are COLOMBO, BREYERS, DANNON, KEMPS, OLD HOME, STONYFIELD FARM, WELLS DAIRY, YOFARM and YOPLAIT.

[3] Households in the sample are assumed to make vendor choices at each shopping trip. For multiple vendor brands on a shopping trip, we use the most heavily chosen vendor.

<Table 1 about Here>

For each household in the yoghurt market, a switch-per-choice (SPC) measure is defined by dividing the number of total switches over the year by the household's total number of choices. A unit value of SPC suggests the household switch brands after every single visit, while the most persistent households have SPC equal to 0. Given a certain level of total visits in a year, a higher SPC indicates more frequent switches. The histogram of SPC (Figure 1) shows that this measure is widely spread between 0 and 1. If conditional on household demographics, SPC explains the state dependence in the targeted carbonated beverage market, we conjecture that those SPC-driven variations reflect heterogeneity in structural state dependence.

<Figure 1 about Here>

Model and Estimation Results

Based on the previous descriptive evidence, we construct unbalanced panel data sets of binary choices for each major vendor in IRI carbonated beverage market in 2007 and match the households with switch information using the IRI yoghurt data sets (Year 2001 and 2002). Denote $Y_{it} = 1$ if household i chooses the targeted brand in shopping trip t, and the model can be written as

$$Y_{it} = 1 \text{ if } \alpha + \beta_0 \text{Price}_0 + \beta_1 \text{Price}_1 + \cdots$$

$$\ldots + [I(Y\_(it-1) = 1)] \# [\text{SPC}_i, \text{Total}_i, \text{Demo}_i]\beta + \epsilon_{it} > 0,$$

where $I(Y_{it-1} = 1)$ represents the state dependence variable, SPC_i represents per-choice switches of each household in the reference category (yoghurt markets), Total_i is the total number of visits in the reference category, and Demo_i consists of household characteristics including household size, age, income, and education. Price_1 is the

normalized transaction price (Dollars per 144 oz of the relevant products) of the targeted vendor brand, and Price_0 indicates the price index for the rest of the products. The operator '#' represents a full interaction between each left and right side variables in the square bracket along with their main effects.

<Table 2 about Here>

Table 2 reports estimation results on a set of random effect logit models. While the main effects of past choice are positive for all brands, the interactive effect with SPC remain negative. The magnitude of the effect is strong, compared with other demographic variables. Two other variables also have consistent interactive effects with past purchases: the total number of shopping trips in the yoghurt market and the household size. The first variable is positively correlated with state dependence. Households who used to make purchases frequently are more likely to maintain their choices for consecutive shopping trips. The second variable is negatively correlated with state dependence. This also seems reasonable, since a larger household has to satisfy more tastes. The rest interactive effects seem to have undetermined signs across different brands.

To better understand the effect of SPC, we calculate the marginal effects in Table 3. On average, past choices increase repurchase probability by 1%--3% based on different vendor brands. In addition, SPC has significant impacts on past choices of national vendor brands. For example, a persistent household in the yoghurt market is subjected to a 6%--8% increase in repurchase probability with Coca-Cola brand, while a frequent switcher does not reveal any systematic bias in state dependence. Compared with the major national brands, private label brands witness weaker state dependence effects and SPC cannot explain them significantly.

<Table 3 about Here>

Therefore, the marginal effects show that for the major brands in the carbonated beverage market, at least part of the heterogeneity in consumers' state dependence is not market- or product- dependent. Nor can it be explained by household characteristics. We conclude that those explained variations in state dependence reflect preference difference that is intrinsic to consumers.

Conclusion

In this paper, we demonstrate that switch information in the yoghurt data five or six years ago (negatively) explains state dependence in a carbonated beverage market. Those explained variations in state dependence reflect preference that is intrinsic to consumers and should be interpreted as structural. Compared with other demographic variables which are commonly used for segmenting a market, switch information has a stronger and more consistent impact on consumers' state dependence effects. Our method of estimating heterogeneity in structural state dependence can be applied to larger and more general field data sets, in which confounding factors of state dependence exist and simulation-based methods are computationally more expensive.

References

Bronnenberg, Bart J., Michael W. Kruger, and Carl F. Mela. 2008. "Database paper: The IRI marketing data set." *Marketing Science* 27(4): 745-748.

Dubé J-P, G.J. Hitsch, P.E. Rossi. 2010. "State Dependence and Alternative Explanations for Consumer Inertia," *RAND Journal of Economics* 41(3): 417–445.

Heckman, James J. 1981. "Heterogeneity and State Dependence." *Studies in Labor Markets.* University of Chicago Press. 91-140.

Keane, Michael P. 1997. "Modeling heterogeneity and state dependence in consumer choice behavior." *Journal of Business & Economic Statistics* 15(3): 310-327.

Rossi, P.E., G.M. Allenby, R. E. McCulloch. 2005. *Bayesian Statistics and Marketing.* John Wiley & Sons. West Sussex. England.

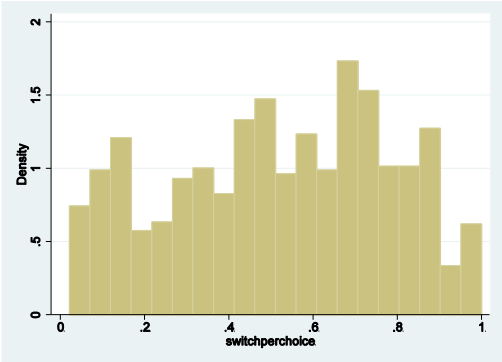
Seetharaman, P. B. 2004. "Modeling multiple sources of state dependence in random utility models: A distributed lag approach." *Marketing Science* 23(2): 263-271.

Train, K.E. 2009. *Discrete Choice Methods with Simulation*, 2nd ed. Cambridge University Press. Cambridge. UK.

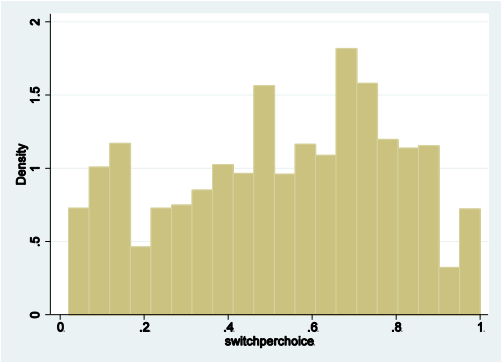
Table 1 Summary Statistics for Households

	Yoghurt MKT				Coke MKT	
	Year 2001		Year 2002		Year 2007	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
No. of Household	3160	-	3783	-	4189	-
Avg Choice: Total	19.1	10	19.7	10.2	15.3	8.6
Avg Switch: Total	8.9	6.6	9.4	7	7.7	8.6

Figure 1 Histogram–Switches Per Choice (Year 2001 and 2002)



Year 2001



Year 2002

Table 2 Estimation Results (Coefficients Shown)

	SPC in Year 2001			SPC in Year 2002		
	Coca Cola	Pepsi	Private	Coca Cola	Pepsi	Private
lagc(hoice)	0.391	0.286	0.429	0.548*	0.0613	0.369
	(0.247)	(0.257)	(0.397)	(0.226)	(0.240)	(0.373)
lagc*SPC	-0.281	-0.302*	-0.779**	-0.407**	-0.105	-0.292
	(0.145)	(0.151)	(0.258)	(0.138)	(0.146)	(0.256)
lagc*total	0.00205	0.0137**	0.00614	0.000270	0.0128**	0.00678
	(0.00419)	(0.00442)	(0.00730)	(0.00362)	(0.00394)	(0.00651)
lagc*hh_size	0.0101	-0.0324	-0.179***	-0.0188	-0.00998	-0.126*
	(0.0302)	(0.0317)	(0.0518)	(0.0289)	(0.0304)	(0.0503)
lagc*hh_age	0.0245	0.000503	0.0201	0.0130	0.0188	-0.0168
	(0.0326)	(0.0351)	(0.0548)	(0.0297)	(0.0319)	(0.0511)
lagc*hh_edu	0.0000897	-0.0345	0.0708	-0.00441	-0.0276	0.0543
	(0.0254)	(0.0271)	(0.0439)	(0.0237)	(0.0253)	(0.0425)
lagc*hh_inc	-0.0337*	0.00517	0.0520*	-0.0216	-0.00273	0.0328
	(0.0144)	(0.0150)	(0.0264)	(0.0139)	(0.0145)	(0.0262)
price1	-1.094***	-1.107***	-0.283**	-1.179***	-1.074***	-0.258**
	(0.0581)	(0.0606)	(0.106)	(0.0543)	(0.0563)	(0.0988)
price0	0.667***	1.208***	1.304***	0.717***	1.213***	1.179***
	(0.0805)	(0.0956)	(0.161)	(0.0744)	(0.0879)	(0.154)
SPC	-0.0448	0.0308	1.263***	0.135	-0.401*	1.306***
	(0.188)	(0.209)	(0.289)	(0.174)	(0.196)	(0.279)
total	-0.0103	-0.0167**	0.0235**	-0.00822	-0.0243***	0.0246***
	(0.00554)	(0.00620)	(0.00821)	(0.00481)	(0.00545)	(0.00743)
hh_size	0.124**	0.114*	0.303***	0.109**	0.149***	0.266***
	(0.0407)	(0.0453)	(0.0599)	(0.0384)	(0.0433)	(0.0590)
hh_age	-0.0528	-0.207***	-0.0715	-0.0654	-0.165***	-0.0982
	(0.0415)	(0.0464)	(0.0620)	(0.0373)	(0.0421)	(0.0583)
hh_edu	-0.0283	-0.0490	-0.0321	-0.00924	-0.0549	0.000441
	(0.0329)	(0.0368)	(0.0494)	(0.0303)	(0.0342)	(0.0473)
hh_income	0.0853***	-0.0172	-0.0961***	0.0838***	-0.0281	-0.0900**
	(0.0191)	(0.0212)	(0.0286)	(0.0179)	(0.0200)	(0.0278)
constant	-0.100	0.615	-6.591***	-0.116	0.639	-6.503***
	(0.381)	(0.415)	(0.610)	(0.343)	(0.375)	(0.574)
week dummies	Yes	Yes	Yes	Yes	Yes	Yes
Cluster	1610	1859	1859	1859	1610	1604
N	25371	28951	28951	28951	25363	24743

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3 Marginal Effects of SPC on State Dependence

	Coca Cola	Pepsi	Private
2001 Yoghurt			
SPC=0	0.0623**	0.0546**	0.0193**
	(0.0196)	(0.0192)	(0.00593)
SPC=0.5	0.0300***	0.0227**	0.0150***
	(0.00848)	(0.00832)	(0.00326)
SPC=1	-0.00118	-0.00812	0.00257
	(0.0165)	(0.0160)	(0.00707)
2002 Yoghurt			
SPC=0	0.0771***	0.0319	0.00991**
	(0.0187)	(0.0197)	(0.00384)
SPC=0.5	0.0315***	0.0188*	0.0127***
	(0.00801)	(0.00802)	(0.00274)
SPC=1	-0.0135	0.00694	0.0141*
	(0.0158)	(0.0143)	(0.00712)

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$