State Dependence and Preference Heterogeneity: The Hand of the Past on Breakfast Cereal Consumption

Dániel Békési[a*], Jens-Peter Loy[b], Christoph Weiss[a]

[a]Department of Economics, Vienna University of Economics and Business, Vienna, Austria
[b]Department of Agricultural Economics, University of Kiel, Germany

Paper prepared for presentation at the 87th Annual Conference of the Agricultural Economics Society, University of Warwick, United Kingdom
8 - 10 April 2013

Abstract

By explicitly accounting for observed and unobserved taste heterogeneity, we investigate state dependence in food consumption on household level. Positive state dependence, which can be interpreted as a loyalty measure, implies habit persistence, whereas negative state dependence implies variety-seeking behavior. Using ConsumerScan data for 2,717 households in Germany for a four-year period (2000 until 2003), we apply mixed multinomial (random coefficient) logit models to analyze consumption behavior in the breakfast cereal market. We find substantial heterogeneity between households: the majority of them express some degree of positive state dependence, which depreciates over time. Merely a small share of the households can be characterized as variety-seekers. The analysis sheds light on the correlation between price sensitivity and loyalty. More state dependent households seem to be more sensitive to price changes. By obtaining household-specific coefficients, we are able to define each household´s position on the distribution of the parameters. Comparing state dependence coefficients and two further loyalty measures (brand-runs and repurchase probabilities), we observe a positive correlation between these, which underpins the importance of the effect of prior consumption. As a diagnostic check, the unconditional distributions of the parameters and the average of the distributions conditional on previous choices are compared.

Keywords Food consumption, state dependence, preference heterogeneity, mixed logit

JEL code Consumer Economics: Empirical Analysis - D12,
Discrete Regression and Qualitative Choice Models - C35

*****Corresponding author
(Vienna University of Economics and Business - Department of Economics,
Augasse 2–6, A – 1090, Vienna, Austria, email: dbekesi@wu.ac.at)

1. Introduction

Studying consumers´ decisions has a long history in social sciences. Making the same choices over time, buying the same product as in previous purchase occasions is also referred as inertia (see early works of e.g. *Brown, 1952*, *Frank, 1962* and *Massy, 1966*). This phenomenon can be explained by two distinct factors. First, prior choices can directly alter current choice probabilities (i.e. true or structural state dependence). If the choice probability conditional on past purchases is higher than the unconditional probability, we refer to positive state dependence, which implies habit persistence, i.e. reinforcement of preferences by past behavior over time (see *Erdem, 1996*). *Dubé et al. (2010)* find that the economic explanation for observed state dependence in consumer choices can be traced back to (brand) loyalty, which has gained an ever growing attention in marketing and economics literature since *Copeland (1923).* For more recent empirical evidence of state dependence see e.g. *Seetharaman et al. (1999)*, *Richards et al. (2007)*, *Arnade et al. (2008)* or *Thunström (2010)*. Negative state dependence, on the other hand, refers to cases where the conditional choice probability is lower than the unconditional probability, implying variety-seeking behavior. In such cases, consumers derive higher utility from switching between different alternatives. A recent survey on the economics of variety-seeking for food products is available in *Weiss (2011)*. Empirical evidence reveals that not accounting for state dependence might yield misleading empirical results *(Erdem, 1996)*.

It is important to note that food consumption behavior exhibits substantial differences between individuals. This leads us to the second possible source of inertia. That is, if decision makers differ in some serially correlated unobserved attributes, which are not appropriately controlled for (see the distinction of *Heckman (1981)* between true and spurious state dependence). Having diabetes for example may lead to repeated choice of non-sugary products, which is, however, not reinforced by previous consumption, but is a result of some consumer background variables and not a genuine behavioral effect. Therefore, in analyzing consumer choice behavior, adequate modeling of unobserved heterogeneity is important.

In this paper, like in many of the above mentioned ones, consumers are assumed to be myopic, or backward-looking. Only the prior consumption path counts and they do not recognize the effect of their current consumption on future preferences; unlike in the specification of *Becker et al. (1994)* applied for smoking or more recently by *Richards et al. (2007)* for snacking.

Early studies like *Manser (1976)* or *Pashardes (1986)* find evidence for habit persistence in aggregate data. However, as *Thunström (2010)* points out, aggregating over different food categories might mask the true nature of past consumption effects. Therefore using detailed micro-level data is advocated, like *Erdem (1996), Seetharaman et al. (1999), Richards et al., 2007*, *Arnade et al. (2008), Dubé et al. (2010)* or *Thunström (2010)* do for categories like: margarine; ketchup, peanut butter, stick margarine, canned tuna; diverse snacks; cheese; refrigerated orange juice, margarine; and breakfast cereals, respectively. These studies find significantly positive state dependence in food consumption by applying discrete choice models.

Using ConsumerScan data for 2,717 households in Germany for a four-year period (2000 until 2003), we apply mixed multinomial (random coefficient) logit models to analyze consumption behavior in the breakfast cereal market. Breakfast cereals are an interesting product group, since there is a huge variety of products in this category to choose from with a considerable variance in prices; and breakfast cereals are also likely to be purchased quite frequently. Furthermore, breakfast choice is of great importance considering health conditions, such as for instance blood sugar level and nutrition in general, as discussed by e.g. *Liljeberg et al. (1999)* and *Nilsson et al. (2008).* Thus, accumulating knowledge about consumer nutritional preferences and habitual behavior points far beyond mere sales considerations.

Our contribution to the body of research of consumer behavior is two-fold. First, we extend and combine different approaches of previous works. Like most of the before mentioned papers, we account for unobserved preference heterogeneity, and as *Seetharaman et al. (1999)* or *Arnade et al. (2008)*, model the depreciation of habit effects over time. By estimating household-specific coefficients, we are also able to define the likely position of a decision maker on the distribution of sensitivities, which can be of great asset and might yield valuable knowledge for market-researchers or policy-makers (see *Hess, 2010*). We compare the estimated-household specific state dependence coefficients with further loyalty measures, namely with the average length of brand runs and repurchase probabilities (see *Mellens et al., 1996*).

Furthermore, there seems to be no unified evidence in the marketing literature about the sign of the correlation between the strength of habit effects (loyalty) and the price sensitivity of households. In case of multiple categories (tuna, margarine, ketchup, peanut butter), *Seetharaman et al. (1999)* find that households with a higher degree of state dependence are less price sensitive, whereas *Erdem and Sun (2002)* report that more use-sensitive ones are also more sensitive to price changes for tooth paste purchases. We investigate this relationship in breakfast cereal consumption for a relatively large set of respondents. Secondly, we perform a diagnostic check of the mixed logit model, suggested by *Train (2009)* and practiced for instance by *Wang et al. (2011)*, which is mostly omitted in similar works. Given that the average of the distributions conditional on prior consumer choices is similar to the unconditional distribution of parameters, the model can be regarded as well specified and accurately estimated.

The present paper is organized as follows. *Section 2* describes the data; *Section 3* presents the estimation method; *Section 4* discusses the empirical model and the estimation results, finally, *Section 5* concludes.

2. Data

The data sample used for the estimation describes the income level and family status of households, the decision maker's age and occupational status for 2,717

German households between 06.01.2000 and 27.12.2003.[1] Households are selected on the criterion that they buy only one kind of a product and brand (e.g. Kellogg´s Honey Loops) on the same purchase occasion. Eventually, there are totally 17,515 observations of 10 brands (alternatives) and 184 different products in the estimation sample. A summary of the demographic variables is found in *Table 1*.

	MAX	*MIN*	*MEAN*	*% of households < mean*	*% of households> mean*
INCOME (€)	*4125*	*250*	*2008.7*	*52*	*48*
AGE (years)	*72*	*19*	*41.7*	*53*	*47*
			% of households		
STRUCTURE				*OCCUPATION*	
singles	*15*			*blue-collar*	*20.3*
w/o children	*34*			*white-collar*	*79.7*
w/ children	*51*				

Table 1: Summary of household characteristics data

In order to control for observed preference heterogeneity, variables for household characteristics were included in the empirical model in *Section 4*, where some modifications of the household data were undertaken. First, "*Income*" is the log of the household´s net monthly income. Since only income groups are given in the data set, the average of the lowest and highest value of a given group is associated to every household in that particular group. Second, "*Age*" stands for the age of the decision maker of the household, again the log of age is taken. Third, "*Structure*" represents the family structure of the household; there are 3 groups: singles, families without children and families with children. Fourth, the "*Occupation*" dummy distinguishes between farmers and workers as well as white-collar civil servants and liberal professionals. The omitted (baseline) categories are: households without children and farmers or blue-collar workers. In the estimation procedure, since prices for the alternatives that were actually not chosen on a particular choice occasion are not available, these were proxied by the weekly average prices of the given brand using the original data set.

Every household has ten mutually exclusive and jointly exhaustive alternatives to choose from - the ten largest breakfast cereals brands in the original data set in terms of sales revenue: i.e. 4 national brands (Kellogg´s, Nestlé, Hahne, Vitalkost) and 6 store brands (Aldi, Lidl, Rewe, Tengelmann, Metro, Edeka). The share of brands chosen by the different household categories is summarized in *Table 2* for the final set of respondents.

[1] The data come from the 'Gesellschaft für Konsumforschung (GfK) – ConsumerScan' database. We are grateful to Janine Empen (Department of Agricultural Economics, University of Kiel, Germany) for preparing the initial data set.

SALES	Kel	Nes	Hah	Vit	Ald	Lid	Rew	Ten	Met	Ede
all	46.58	8.08	1.64	1.70	24.33	5.91	5.43	3.19	1.66	1.47
income (€)										
$<$ *ave (2009)*	42.91	7.80	2.20	2.37	24.36	6.70	6.61	3.82	1.81	1.41
$>$ *average*	50.10	8.35	1.11	1.06	24.30	5.16	4.31	2.59	1.50	1.52
age (years)										
$<$ *ave (41.7)*	43.27	8.46	1.58	2.10	25.00	6.59	5.56	3.54	2.36	1.56
$>$ *average*	49.48	7.75	1.70	1.35	23.75	5.32	5.33	2.89	1.04	1.39
structure										
singles	40.34	8.53	1.39	2.25	25.66	4.98	8.15	5.16	1.82	1.71
w/o child	53.92	7.58	2.40	1.45	19.65	5.02	5.10	2.22	1.39	1.28
w/ child	43.80	8.27	1.27	1.72	26.70	6.63	5.02	3.30	1.77	1.52
occupation										
blue-collar	39.88	10.23	1.24	1.94	27.60	6.72	5.53	3.86	1.31	1.68
white-collar	48.41	7.50	1.75	1.64	23.44	5.70	5.41	3.01	1.75	1.41

Table 2: Relative market shares in % by household groups and brands (Columns for brands)

3. Estimation Method

Assume that the utility function of individual *n* choosing alternative *j* at time period *t* is:

$$U_{njt} \quad = \quad \beta_n' x_{njt} + \varepsilon_{njt} = (\bar{\beta}' + \xi_n')x_{njt} + \varepsilon_{njt} \tag{1}$$

where β_n is specific for every household; it has a mean (common) part and an household-specific part, formally: $\beta_n = \bar{\beta} + \xi_n$. The error term ($\varepsilon$) is Gumbel independently and identically distributed (i.i.d.), and β is unknown. Following *Train (2009)*, if we knew β_n, the probability of household's *n* observed choice sequence $y_n = \{y_{n1}, \ldots, y_{nT}\}$ from $j = 1, \cdots, J$ alternatives described by some variables *x* would be:

$$P\left(y_n \mid x_n, \beta\right) \; = \prod_{t=1}^{T} \frac{e^{\beta' x_{n y_{nt} t}}}{\sum_j e^{\beta' x_{njt}}} \tag{2}$$

As β is unknown to the researcher, the probability of household's *n* sequence of choices is the integral of *Eq. 2* over the distribution of β.

$$P\left(y_n \mid x_n, b, W\right) \; = \int\limits_{\beta} P\left(y_n \mid x_n, \beta\right) \phi(\beta \mid b, W) d\beta \tag{3}$$

This mixed logit formula is a weighted average of the standard logit probability calculated at different values of β. The weight is the probability density (ϕ) of β in the entire population of households, with mean *b* and the variance-covariance matrix W, assuming normal distribution (see *Train, 2009*). Mixed logit models - MXL were used first in the transportation literature on aggregated vehicle choice data (see e.g. *Boyd and Mellman, 1980*). More recent works that use mixed logit models and do research on micro-level breakfast cereal choice data are e.g. *Nevo (2001)* - who accounts for unobserved heterogeneity but not for habit persistence - or *Thunström (2010),* who takes both into consideration.

This widely used framework has several advantages over other traditional discrete choice models, like the less flexible standard multinomial logit - MNL model. Firstly, parameters can be defined to be heterogeneous over households, i.e. random taste variation can be accounted for. *Heckman (1981)* distinguishes between true and spurious state dependence. In the first case, preferences that play a role in present choices are affected by previous experiences. ´The conditional probability that an individual will experience the event in the future is a function of past experience.´ If this holds, then there is a ´genuine behavioral effect´. However, if consumers differ in some unmeasured variables that are independent of the past experience, but alter the consumer´s probability to make a certain choice, than the estimation of the effect of previous decisions will not be reliable. Hence, unobserved preference heterogeneity has to be controlled for, i.e. the parameters of the model are defined to be random over households.

A further advantage of the mixed logit model is that the IIA (independence of irrelevant alternatives) assumption does not have to hold; utility correlates over alternatives. The relative probabilities of two alternatives in the mixed logit setup depend on all the other alternatives. MXL allows correlation in unobserved factors also over time (purchase occasions). Therefore, unlike in the standard logit model, unobserved factors that are persistent over time can be accounted for.

The mixed logit (like the standard logit) model has another beneficial feature that makes it particularly suitable for the investigation of habit persistent behavior. The inclusion of a lagged dependent variable, in order to account for state dependence is not problematic. Conditional on β, the only remaining random terms in (*Eq.1*) are the ε_{njt}'s, which are assumed to be i.i.d., independent over time. Thus, these are uncorrelated with the lagged dependent variable in time period *t* (see *Train, 2009*). Eventually, the mean of β (*b*) and the variance-covariance matrix (W) have to be estimated.

There is, however, a major drawback of using a mixed logit model. Since there is no closed formula for the integral in the mixed logit probability, it is very computer-intensive to calculate it; it has to be approximated through simulation.[2] The simulated probability, which is an unbiased estimator of P is:

[2] Computation time can be considerably reduced by using Halton draws instead of random draws. *Train (2000)* finds that the simulation error was smaller with 100 Halton draws than with 1000 random draws, whereas the computation time was decreased as well.

$$\hat{P}\left(y_n \mid x_n, b, W\right) = \frac{1}{R}\sum_{r=1}^{R}\prod_{t=1}^{T}\left[\frac{e^{\beta^{r\prime} x_{n y_{nt} t}}}{\sum_j e^{\beta^{r\prime} x_{njt}}}\right] \tag{4}$$

The simulated log-likelihood function to be maximized wrt. *b* and *W* is:[3]

$$SLL = \sum_{n=1}^{N} ln(\hat{P}\left(y_n \mid x_n, b, W\right)) \tag{5}$$

Following *Train (2009)*, after estimating *b* and *W*, the distribution of β conditional on the observed choice sequence of household *n* can be derived.

$$\theta\left(\beta \mid y_n, x_n, b, W\right) = \frac{P\left(y_n \mid x_n, \beta\right)\phi(\beta \mid b, W)}{P\left(y_n \mid x_n, b, W\right)} \tag{6}$$

The mean of this distribution conditional on the observed choices of households that choose y_n given x_n can also be obtained and considered as an estimate of β_n. Hence, it reveals the preference of a particular household given its prior choices. (See *Train, 2009* for further details):

$$\bar{b}_n = \frac{\int_{\beta} \beta \cdot P\left(y_n \mid x_n, \beta\right)\phi(\beta \mid b, W)d\beta}{P\left(y_n \mid x_n, b, W\right)} \tag{7}$$

If the average of the conditional taste distributions (θ) across households is similar to the estimated population (unconditional) distribution (ϕ), than it can be concluded that the model specification is correct and the estimation is accurate. This diagnostic check is performed in *Section 4.2*.

4. Empirical Specification and Estimation Results[4]

4.1. Identification and parameters of the population distribution

In discrete choice models, the dependent variable is a dummy which is set equal to 1 if the alternative was indeed chosen and zero otherwise. Following *Thunström (2010)* and others, household characteristics are included in the model in order to be able to control for observed heterogeneity of preferences. Note that household variables do not vary within a choice occasion. Therefore, these must be interacted

[3] It must be noted that because of the non-linearity of the log transformation, $ln\hat{P}$ is not an unbiased estimator of lnP, hence the MSLE (maximum simulated likelihood estimator) is also biased. But fortunately, this bias reduces, if the number of draws (i.e. *R*) increases and eventually, the estimator is consistent and equivalent to the classical maximum likelihood estimator (see *Train, 2000*). For the estimation we used 100 Halton draws.

[4] The estimations of the means and the variance-covariance matrix were conducted in Stata 12 with the user-written module of Arne Risa Hole called -mixlogit- (see *Hole, 2007*).

with the alternative-specific constants (see *Hole, 2007*). Estimates of demographic effects must be interpreted as effects relative to a reference alternative (brand) like in the multinomial logit model. To account for the observed heterogeneity in state dependence, we also include interaction terms between state dependence and household characteristics variables. For the sake of traceability, these interaction variables are not presented in *Eq. 8*. The unobserved part of preference heterogeneity is captured by the random parameters (price and state dependence), which are allowed to vary across households. The utility of household *n* from purchasing brand *j* on purchase occasion *t* is:

$$\begin{gathered} U_{njt} = \alpha_j + [\beta'_{sdep,n} + \beta'_{WearOut} Ln(T)_t] StateDep_{jt} + \beta'_{Price,n} Price_{jt} \quad (8) \\ + \beta'_{Income,j} Ln(Income)_j + \beta'_{Age,j} Ln(Age)_j \\ + \sum_{k=1}^{3} \beta'_{Family,j,k} Family_{j,k} + \beta'_{Occupation,j} Occupation_j + \varepsilon_{njt} \\ t = 1...T;\ n = 1...2717;\ j = Kelloggs...Edeka \\ k = single,\ HH\, w/childern,\ HH\, w/o\, children \end{gathered}$$

where α_j is the alternative-specific constant of alternative *j*. Aldi is defined to be the reference category; the coefficients of household-specific variables that are reported express effects relative to this benchmark alternative. Following *Thunström (2010)*, state dependence in consumption behavior is introduced by the one period lagged dependent variable. The indicator variable $StateDep$ is set equal to 1 if the given brand was purchased a period before and zero otherwise. If the estimated coefficient is positive, households are characterized by habit persistence; a negative parameter estimate indicates variety-seeking behavior. The variable *Price* is defined to be random in order to account for possible heterogeneity in price sensitivity across households. We estimate generic parameters for *StateDep* and *Price* for each brand in order to keep the model computable. Wald tests conducted after a multinomial logit estimation cannot justify that all parameters should be considered being different. Moreover, ten times two alternative-specific random parameters would most obviously lead to estimation difficulties and non-convergence. Furthermore, we allow these parameters to be correlated with each other. In order to capture the time effect, we allow state dependence to vary over time. *WearOut* is an interaction of the log of time in days between the current and the last purchase occasion and state dependence. If the estimated coefficients are negative, we can conclude that the effects of previous purchases depreciate over time (see *Seetharaman et al., 1999*). The parameters of the household demographic variables, the alternative specific constants and the above-mentioned interaction effects are fixed (like in the standard logit model), i.e constant across households. The error term is Gumbel i.i.d.

The standard multinomial logit model - *MNL* is reported in *Table 4*. The difference between *MXL* and *MXL+inter* is that the interaction effects with state dependence are included only in the later one. Other than that the two models are identical. Note that the first two parameters of these are allowed to differ between households and only the more interesting estimates are presented. Comparing MNL and MXL specifications, we can conclude that not controlling for preference heterogeneity can lead to overestimated state dependence. It is clear from the likelihood-ratio tests that the MXL models with the correlated random parameters are superior to a model where all the parameters are fixed (MNL). The standard deviations (the diagonal elements of the variance-covariance-matrix, W) are statistically significant, indicating that the random specification of the given parameters is correct (see *Thunström, 2010*). Ommiting an alternative and using the Hausmann-test, it is clear that the IIA assumption is violated in the standard MNL model. According to a final LR-test *MXL+inter* fits the data better than *MXL* or a model version, where the parameters are not correlated, i.e W is diagonal. The off-diagonal elements of (W) are also statistically significant, emphasizing that defining the random parameters to be correlated is correct. That taken together, applying the mixed logit framework with correlated parameters seems to be justified.

The results are quite similar across these model specifications. There are only two major differences. The price coefficient in the *MNL* Model is negative, but insignificant; and the estimate of state dependence is quite different in the last two models. Controlling for interactions with household characteristics may explain the later finding. There seems to be positive state dependence on average in the consumption of all cereal brands (indicating habit persistence). For the average household the marginal utility effect of choosing a cereal brand in the previous period is positive on choosing that specific brand again. Nevertheless, a really small portion of households are variety-seekers (negative state dependence); they are in a clear minority, 8% of the total population in *MXL* and and 0.4% if we account for observed household heterogeneity in state dependence.[5] These effects depreciate over time, as the estimated coefficient of *WareOut* is negative.

With regards to the impact of household characteristics on state dependence, we find that higher income households tend to be less habitual compared to households with a lower level of income. We find the same for households with children, which most probably could be explained by changing tastes of children or by the bigger size of these households. On the other hand, older decision-makers are more affected by prior choices. The price coefficient is negative on average for the majority of households, but some of them (27%) seem to associate some kind of quality to higher prices, as they are more likely to purchase a brand if the price is higher. A similar finding is reported by *Nevo (2001)*. Taking a look at predicted choice probabilities, a ceteris paribus 25% increase in the price (€/100g) of a given brand has the largest effect averaged across households (-6.6%) in case of Aldi. Vitalkost is the least affected with only -1%. The average effect across brands is -2.3%. Since we have not attempted to estimate ten different price coefficients in the mixed logit framework, these findings are rather valuable for distinguishing reactions of

[5] Calculated from a cumulative normal distribution table. The cumulative distribution functions (CDF) are presented in *Subsection 4.2*.

Table 3: Model estimates

	Stand. MNL	MXL		MXL+inter	
	MEAN	MEAN	SD	MEAN	SD
Price	-0.016	-6.574***	10.750***	-6.536***	10.553***
	(-0.10)	(-17.17)	(26.06)	(-17.37)	(26.10)
StateDep	4.322***	2.255***	1.625***	3.997***	1.542***
	(10.67)	(44.92)	(27.92)	(4.54)	(29.01)
incomeRewe	-0.463***	-0.400***		-0.416***	
	(-4.20)	(-3.08)		(-3.18)	
incomeTengelmann	-0.429***	-0.479***		-0.501***	
	(-3.60)	(-3.51)		(-3.65)	
ageMetro	-1.193***	-1.399***		-1.398***	
	(-5.72)	(-5.86)		(-5.82)	
singleHahne	-1.047***	-0.759***		-0.800***	
	(-4.76)	(-2.98)		(-3.11)	
childKellogg	-0.344***	-0.923***		-0.818***	
	(-4.63)	(-6.08)		(-5.34)	
childNestle	-0.208**	-0.547***		-0.503***	
	(-2.06)	(-2.96)		(-2.73)	
occupKellogg	0.141**	0.498***		0.529***	
	(2.01)	(3.31)		(3.58)	
WearOut	-0.335***			-0.278***	
	(-22.36)			(-13.17)	
State Dep.age	0.419***			0.287**	
	(6.43)			(1.97)	
State Dep.income	-0.248***			-0.192*	
	(-5.15)			(-1.86)	
State Dep.single	-0.221***			0.160	
	(-3.18)			(1.03)	
State Dep.child	-0.437***			-0.460***	
	(-9.33)			(-4.43)	

z scores in parentheses
*p<0.10 **p<0.05 ***p<0.01
MNL LL = -19585.021
MXL LL = -17610.302
MXL+inter LL = -17505.527

consumers. Buyers are more price sensitive in case of of store brands compared to national brands (-2.6% vs. -1.9%). Higher income households, households with older decision-makers, singles and households with children are found to purchase Rewe, Tengelmann, Metro, Hahne, Kellogg´s, Nestlé less probably, respectively compared to their counterparts relative to the reference group (Aldi).

In the variance-covariance matrix (W) of *Model MXL*[6] in *Table* 4, we observe negative and statistically significant covariance between the coefficients of *StateDep* and *Price*, i.e. price sensitivity gets higher, if state dependence increases.

	β_{Price}	$\beta_{StateDep}$
β_{Price}	*115.56*** (1)*	*-4.16*** (-0.24)*
$\beta_{StateDep}$		*2.64*** (1)*

Table 4: Variance-covariance matrix (W), Correlation coefficients in parentheses

At the first sight, this may seem a bit surprising, as conventional wisdom would suggest that more state dependent decision-makers are also less price-sensitive. This is indeed found by e.g. *Seetharaman et al. (1999).* Although this seems to be much more like an empirical rather than a theoretical question, as the findings are far from conclusive and our result cannot simply be labeled as counter-intuitive either. In various fields our finding is emulated. *Allenby and Lenk (1995)* (ketchup), *Reinartz and Kumar (2002)* (catalog retailer), *Erdem and Sun (2002)* (tooth paste) or *Petrick (2005)* (cruise ships) suggest that consumers with more frequent purchases, higher use-sensitivity or attachment to a specific alternative may indeed become more price-sensitive than occasional ones.

This argumentation might be explained as follows. Frequent (or loyal) consumers are probably more aware of the quality, value and past prices of a given product and/or other alternatives. Repetitive buying or strong preference for a given brand could increase consumer's recall of brand's price and hence the importance of a consumer's reference price *(Rajendran and Tellis, 1994).* This is an internal benchmark influenced by past and alternative prices which is compared to actual prices of other alternatives. According to *Thomas and Menon (2007)*, more confident consumers tend to have a lower reference price, thus they are more price-sensitive if price plays a substantial role in the decision-making process. They also might have a smaller consideration set *(Erdem and Sun, 2002)* and therefore can make comparisons more easily, which might increase their price sensitivity. Since breakfast cereals are close substitutes, switching cannot be difficult. Alternatively, this phenomenon might also be explained by search costs. As consumers can save more money on frequently purchased products, searching cheaper alternatives becomes more beneficial, hence price sensitivity may be higher (see e.g. *Sorensen, 2000*). The author finds that due to the expected benefits of search, prices of frequently purchased prescriptions exhibit reductions in price dispersion and price-cost margins.

[6]We use this model specifications instead of the *MXL+inter* as the coefficients of state dependence here incorporate heterogeneity otherwise captured partly by the interaction terms.

If consumers are rather influenced by the price of a product and not so much by other attributes of a brand, than those who otherwise would stick to a specific brand with an unchanged low price, are going to react sensitively as the price increases. They are actually loyal to the price of an alternative and not so much to the brand itself (*Brown, 1953*).

We have to stress, however, that in this case, we have not distinguished different brands; we estimated generic state dependence and price parameters for all the alternatives. This might be seen as a limitation of the mixed logit framework, if the researcher uses a data set with a high number of alternatives, even if the restriction seems to be justified in this case.

4.2. Individual-specific estimates and diagnostic check

After *b* and *W* are estimated in *Model MXL*, one can obtain household-specific parameter estimates ($\bar{b}_n$), as defined in *Eq. 7*.[7] *Figure 1* depicts the correlation between the estimated household-specific coefficients of state dependence and price. The negative relationship is quite apparent (correlation: -0.28), which underpins the results in *Table 4* - the variance-covariance matrix from the population (unconditional) distribution. The more positive price coefficient is (interpret higher prices as a measure of quality), the stronger variety-seeking behavior households exhibit, which makes sense intuitively.

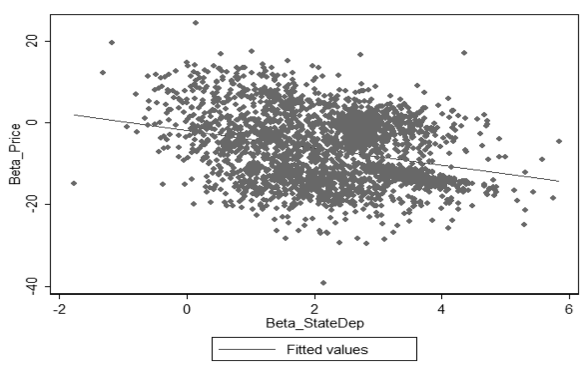

Figure 1: Household-specific coefficients, scatter-plot of the conditional means

[7]The Stata post-estimation command -mixlbeta- of *Hole (2007)* was used to obtain the means of conditional distributions for every household.

Figure 2 represents the cumulative distribution functions (CDF) of these means of the conditional distributions across households (i.e. the most likely values of the coefficients for each respondent) compared to the normal CDFs of the parameter estimates - with the means and standard deviations estimated in *Subsection 4.1*. These reinforce our above findings. Indeed only a minor proportion of households are variety-seekers (see vertical line at zero); in the meantime the distribution of the conditional means has a narrower range for both parameters. In case of the distribution of the conditional means, the portion of variety-seekers is a bit lower. Other than that the CDFs look quite similar.

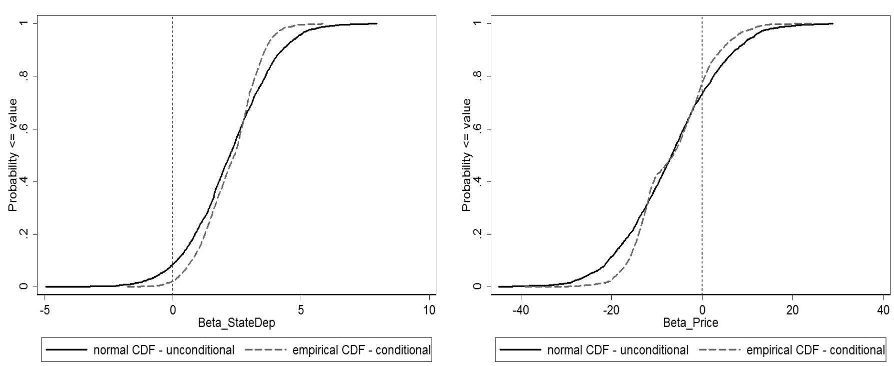

Figure 2: Unconditional distributions and distribution of conditional means

Table 5 compares the estimated means - *b* and standard deviations - sd_β *(i.e.* square roots of the diagonal elements of *W)* of the parameters in the population (unconditional) distribution with the means of the distribution of the conditional means - $\mu_{\bar{b}_n}$ and the standard deviations of the conditional means - $sd_{\bar{b}_n}$. The close similarity of the means indicate a well-specified and accurately estimated model. From *Table 6* it is evident that variation in $\bar{b}_n$ captures almost 70 and 80 percent of the total estimated variation in the coefficients of *Price* and *StateDep*, respectively. This result implies that the mean of a customer's conditional distribution is potentially able to distinguish decision-makers in a meaningful way (see *Train, 2009*). However, these individual-specific coefficients are not known with certainty; in that case sd_β and $sd_{\bar{b}_n}$ would be equal *(Hess, 2010)*.

	b	$\mu_{\bar{b}_n}$	$\frac{\mu_{\bar{b}_n}}{b}$	sd_β	$sd_{\bar{b}_n}$	$\frac{sd_{\bar{b}_n}}{sd_\beta}$
$\beta_{StateDep}$	*2.25*	*2.26*	*1.00*	*1.63*	*1.11*	*0,68*
β_{Price}	*-6.57*	*-6.60*	*1.01*	*10.75*	*8.23*	*0,77*

Table 5: Means and sd for unconditional distributions and for distributions of cond. means

4.3. Alternative-specific loyalty measures

Finally, two further measures of (brand) loyalty were calculated for all households and all brands individually: the average length of brand runs and the repurchase probabilities.[8] Note that these descriptive statistics, other than state dependence in our estimated model, are calculated for all the alternatives, and as a result, we have both measures for all the ten brands and for all the 2,717 households.[9] Brand runs are defined as sequences of consecutive purchases of the same brand. Repurchase probabilities are calculated from switching matrices. A switching matrix simply shows in this case for each household the number of switches between brands and the number of purchase occasions buying the same brand consequently.

	BR_k	BR_n	BR_h	BR_v	BR_a	BR_l	BR_r	BR_t	BR_m	BR_e
S.Dep.	0.51	0.33	0.52	0.53	0.60	0.48	0.61	0.55	0.42	0.58
	RP_k	RP_n	RP_h	RP_v	RP_a	RP_l	RP_r	RP_t	RP_m	RP_e
S.Dep.	0.63	0.52	0.40	0.55	0.54	0.52	0.61	0.57	0.44	0.59
	PF_k	PF_n	PF_h	PF_v	PF_a	PF_l	PF_r	PF_t	PF_m	PF_e
S.Dep.	0.23	0.19	0.35	0.42	0.28	0.22	0.31	0.26	0.32	0.49

Table 6: Correlations of loyalty measures. BR: brand run, PR: repurchase probability, PF: purchase frequency, indexes: brands

As presented in *Table* 6, quite a strong correlation (for the 10 brands ranging from 0.33 to 0.63) was found between these descriptive statistics and our household-specific (but generic across brands) state dependence estimates - obtained in *Eq.* 7. This result may indicate that accounting for heterogeneity in habit persistence might have been successful and different households were rightfully distinguished in terms of their sensitivity to previous consumption and habits in breakfast cereal consumption. It can be concluded that these measures, however obtained by very different methods, show similar results and reinforce the importance of state dependent behavior. According to *Frank (1962)*, habitual purchasing behavior is ´defined as the existence of positive association between the number of previous purchases (frequency) of a particular brand and the consumer repurchase probability´. This is in line with our findings as well (correlation coefficients range from 0.37 to 0.68 for the ten brands - not presented in the table). State dependence is also positively correlated with purchase frequency.

[8]For a detailed discussion of these and other loyalty measures, the reader is referred to *Mellens et al. (1996)*.

[9]We are grateful to Janine Empen (Department of Agricultural Economics, University of Kiel) for the assistance in the calculations of these alternative measures.

5. Conclusions

As significant effects of consumer habit persistence was found, policy interventions or marketing campaigns might have a lagged effect on the targeted population group, because habits may prevent them from reacting to these promptly. This effect gets weaker over time. Therefore, outcomes in the short run could differ from outcomes in the long run. These effects of past consumption are particularly strong for older decision-makers and weaker for higher income households. From examining individual level household-specific sensitivities, these findings about household heterogeneity and state dependence are reinforced and it is also underpinned that a relevant proportion of households might perceive higher prices as a measure of quality. For a significant number of households, a price reduction of cereals will not increase the probability of choosing this product. Again, the effects of price changes (increases or decreases) must be carefully considered and different groups of consumers (in regard of their price sensitivity and habitual behavior) need to be distinguished to enhance the effectiveness of a possible intervention.

Households that are more dependent on prior consumption are found to be more price-sensitive. An explanation for this finding might be that most consumers focus on a specific attribute of the brands, namely price. They become loyal to that given brand, as long as the price remains unchanged, but switch to another alternative, as price increases. Higher search costs for frequently purchased products might also explain this phenomenon, as consumers can save more money on frequently purchased products, therefore searching cheaper alternatives is more beneficial, hence price sensitivity may be higher. Frequent and confident buyers may compare prices more easily and have a lower reference price, which might lead to higher price sensitivity. State dependence is found to be positively correlated with purchase frequency. The main message here for further research is stressing the importance to clarify the source of brand loyalty, i.e. the attribute of a brand the decision-maker is attached to, as this might have vital implication on pricing strategies.

An important restriction of our analysis, however, need to be kept in mind while interpreting its findings. The assumption that households buy only one alternative on a purchase occasion leads to the exclusion of a large number of observations. This is, however, a necessary restriction of discrete choice models. An alternative solution to this limitation is presented e.g. by *Bhat (2005)*, introducing a multiple discrete-continuous version of the multinomial logit model, however, at the cost of a more complicated model formulation.

An approach of examining the relationship between habit persistence and price sensitivity not only on a brand-group level but on a nutrient basis, like calorie, or sugar content (see *Richards et al., 2007*), could also contribute to a better understanding of the nature of nutrition choices made by various individuals. It would also touch upon relevant issues like health-care, medical spending or food safety. Despite the significant heterogeneity in consumers' preferences and food consumption behavior, future research in this area will hopefully improve our ability to derive some general observations on the ´hand of the past´ that might shed more light on the underlying decision-making process of individuals.

References

Allenby, G. and Lenk, P.: 1995, *Brand Loyalty, Price Sensitivity, and Merchandising Effects on Consumer Brand Choice.*, Journal of Business and Economic Statistics, Vol. 13, No. 3: 281-289.

Arnade, C., Gopinath, M. and Pick, D.: 2008, *Brand Inertia in U.S. Household Cheese Consumption.*, American Journal of Agricultural Econommics, 90(3): 813-826.

Becker, G. S., Grossman, M. and Murphy, K.: 1994, *An Empirical Analysis of Cigarette Addiction.*, American Economic Review, Vol. 84, No. 3: 396-418.

Bhat, R. C.: 2005, *A multiple discrete-continuous extreme value model: formulation and application to discretionary time-use decisions.*, Transportation Research, Vol.39B: 679-707.

Boyd, J. H. and Mellman, E.: 1980, *The Effect of Fuel Economy Standards on the U.S. Automotive Market: An Hedonic Demand Analysis.*, Transportation Research, Vol.14A: 367-378.

Brown, G. H.: 1953, *Brand Loyalty - Fact or Fiction.*, Trademark Reporter Vol. 43: 251-258.

Brown, T. M.: 1952, *Habit Persistence and Lags in Consumer Behaviour.*, Econometrica, Vol. 20, No. 3: 355-371.

Copeland, M. T.: 1923, *Relation of Consumers' Buying Habits to Marketing Methods.*, Harvard Business Review 1, 2, 282-289.

Dubé, J.-P., Hitsch, G. J. and Rossi, P. E.: 2010, *State dependence and alternative explanations for consumer inertia.*, The RAND Journal of Economics, Volume 41, Issue 3, 417-445.

Erdem, T.: 1996, *A Dynamic Analysis of Market Structure Based on Panel Data.*, Marketing Science, Vol. 15, No.4: 359-378.

Erdem, T. and Sun, B.: 2002, *An Empirical Investigation of the Spillover Effects of Advertising and Sales Promotions in Umbrella Branding.*, Journal of Marketing Research, Vol. 39, November: 408-420.

Frank, R. E.: 1962, *Brand Choice as a Probability Process.*, The Journal of Business, 35(1): 43-56.

Heckman, J. J.: 1981, *Heterogeneity and State Dependence. [in:] Rosen, S. (ed.), Studies in Labor Markets: 91-140. National Bureau of Economic Research*, McGraw-Hill Companies.

Hess, S.: 2010, *Conditional parameter estimates from Mixed Logit models: distributional assumptions and a free software tool.*, Journal of Choice Modelling, 3(2): 134-152.

Hole, A. R.: 2007, *Fitting mixed logit models by using maximum simulated likelihood.*, Stata Journal, Vol. 7(3): 388-401.

Liljeberg, H. G., Åkerberg, A. K. and Björck, I. M.: 1999, *Effect of the glycemic index and content of indigestible carbohydrates of cereal-based breakfast meals on glucose tolerance at lunch in healthy subjects.*, American Journal of Clinical Nutrition, 69: 647-55.

Manser, M. E.: 1976, *Elasticities of Demand for Food: An Analysis Using Non-Additive Utility Functions Allowing for Habit Formation.*, Southern Economic Journal, Vol. 43, No. 1: 879-891.

Massy, W. F.: 1966, *Order and Homogeneity of Family Specific Brand-Switching Processes.*, Journal of Marketing Research, 3(1): 48-54.

Mellens, M., Dekimpe, M. G. and Steenkamp, J. B. E. M.: 1996, *A Review of Brand-Loyalty Measures in Marketing.*, Tijdschrift voor Economie en Management, Vol. XLI, 4: 507-533.

Nevo, A.: 2001, *Measuring Market Power in the Ready-to-Eat Cereal Industry.*, Econometrica, Vol. 69, No. 2, 307-342.

Nilsson, A. C., Östman, E. M., Granfeldt, Y. and Björck, I. M.: 2008, *Effect of cereal test breakfasts differing in glycemic index and content of indigestible carbohydrates on daylong glucose tolerance in healthy subjects.*, American Journal of Clinical Nutrition, 87, 645-654.

Pashardes, P.: 1986, *Myopic and Forward Looking Behavior in a Dynamic Demand System.*, International Economic Review, Vol. 27, No. 2: 387-397.

Petrick, J. F.: 2005, *Segmenting cruise passengers with price sensitivity.*, Tourism Management, 26: 753-762.

Rajendran, K. N. and Tellis, G. J.: 1994, *Contextual and Temporal Components of Reference Price.*, Journal of Marketing Vol. 58, No. 1: 22-34.

Reinartz, W. and Kumar, V.: 2002, *On the Profitability of Long-Life Customers in a Noncontractual Setting: An Empirical Investigation and Implications for Marketing.*, Journal of Marketing Vol. 64, No. 4: 17-35.

Richards, T. J., Petterson, P. M. and Tegene, A.: 2007, *Obesity and Nutrient Consumption: A Rational Addiction?*, Contemporary Economic Policy, Vol. 25, No. 3: 309-324.

Seetharaman, P. B., Ainslie, A. and Chintagunta, P. K.: 1999, *Investigating Household State Dependence Effects across Categories.*, Marketing Science, 36(4): 488-500.

Sorensen, A. T.: 2000, *Equilibrium Price Dispersion in Retail Markets for Prescription Drugs.*, Journal of Political Economy, Vol. 108, No.4: 833-850.

Thomas, M. and Menon, G.: 2007, *Internal Reference Prices and Price Expectations Diverge: The Role of Confidence.*, Journal of Marketing Research, Vol. XLIV: 401-409.

Thunström, L.: 2010, *Preference Heterogeneity and Habit Persistence: The Case of Breakfast Cereal Consumption.*, Journal of Agricultural Economics, Vol. 61, No. 1: 76-96.

Train, K.: 2000, *Halton Sequences for Mixed Logit.*, Working Paper Series 228, Department of Economics, Institute for Business and Economic Research, UC Berkeley.

Train, K.: 2009, *Discrete Choice Methods with Simulation, Second Edition.*, Cambridge University Press.

Wang, J., Hockenberry, J., Chou, S. and Yang, M.: 2011, *Do bad report cards have consequences? Impacts of publicly reported provider quality information on the CABG market in Pennsylvania.*, Journal of Health Economics, Volume 30, Issue 2, March: 392-407.

Weiss, C. R.: 2011, *Consumer Demand for Food Variety.*, [in:] Lusk, J.L., Roosen, J., and Shogren, J.F. (Eds.), The Oxford Handbook of The Economics of Food Consumption and Policy, Chapter 27: 667-694.