TOWARDS HIGH RESOLUTION IDENTIFICATION OF VARIETY-SEEKING BEHAVIOR

Ho, Edward, ETH Zürich, Chair of Information Management, Department of Management, Technology and Economics, Weinbergstrasse 56/58, 8092 Zürich, Switzerland, eho@ethz.ch

Ilic, Alexander, Institute of Technology Management, University of St. Gallen, Dufourstrasse 40a, 9000 St. Gallen, Switzerland, alexander.ilic@unisg.ch

Abstract

Because smartphones are now ubiquitous, it becomes for the first time economically feasible to operationalize personalized marketing measures also for physical grocery retailing. A particularly interesting and high value target group in this domain is the one of variety seeking, since this group is most likely to respond positively to new offers and recommendations. However, present methods in identifying variety seekers rely on questionnaires and ignore that variety seeking may differ between product categories. In this paper, we present a model for measuring variety seeking behavior on a high level of granularity, based on a consumer's purchases in individual product categories. Our study has three main contributions. Firstly, we contribute to the customer segmentation research stream by providing a novel way for identifying customers' overall extent of variety seeking as well as their specific variety seeking at a category level. Second, for the most important retail categories we characterize the extent of variety seeking and provide a data-driven approach that is easy to operationalize by practitioners – especially for deploying large-scale personalized marketing measures in social or mobile commerce in physical stores. Finally, we provide a method to reconcile the highly granular category-level results with existing per person typologies.

Keywords: Retail, e-commerce, Consumer Segmentation, Variety Seeking.

1   Introduction

It is widely accepted in online e-commerce that granular marketing is more effective than applying the same marketing policy to all consumers (Jiang & Tuzhilin 2006; Changchien et al. 2004). Now with the increasing ubiquity of smartphones used in physical stores, it becomes for the first time technologically and economically feasible to operationalize personalized marketing also for physical retailing via mobile recommender systems (RS). The ability to offer the right product from the right category given a consumers' individual preference requires a fine grained understanding of a consumer's buying behavior.

However, today in physical retailing, marketing methods mainly depend on typologies or classification schemes that segment consumers into different groups according to their overall characteristics (Ailawadi et al. 2001; Chandon et al. 2000; Urbany et al. 1996). A particularly interesting and high value target group is the one of variety seeking, since this group is most likely to respond positively to new offers. To identify these customers, the costly method of deploying questionnaires is often used (Garretson et al. 1998; Fisher et al. 2002; Kerin et al. 1992; Wakefield & Inman 2003; Dodds et al. 1991; Voss et al. 2003; Trijp et al. 1996). However, this approach is not scalable; retailers cannot sustainably survey all of their consumers, and thus the standard marketing action is not personalization, but rather having all consumers receive the same promotion, typically a price cut. Furthermore, current questionnaire approaches neglect that the actual disposition of variety seeking of a person might in fact differ from product category to product category.

Our study is thus motivated to address this gap by giving researchers and in-store retailers the ability to finely identify their variety seekers based on in-store purchasing data, and enable RS to use this insight for marketing action. As one of the first studies in this area, we present a model for a higher resolution classification scheme that allows for measuring a consumer's different variety dispositions on a product category level (Section 3). In contrast to other models (reviewed in Section 2), it does not require explicitly stated preferences or motivations and is purely data-driven. It can be easily implemented by physical retailers with already available in-store point-of-sale (PoS) data and would allow for measuring changes in variety seeking dispositions almost in real-time (shown in Section 4).

The key contributions of this paper are that we prove that current typologies neglect critical properties required for personalized marketing – especially in the context of mobile and social commerce. Also, we are one of the first studies to establish a stochastic counting-based model for high resolution variety seeking and validate the model fit with real PoS data of over 350,000 transactions (Section 4). Finally, we provide a method to reconcile the high resolution and data-driven approach with the currently dominant approach of using an aggregated variety typology per customer (Section 4). The result is a parsimonious model of variety seeking that can be estimated on observable consumer actions, rather than latent motivations, using only consumer purchasing data logged at the PoS (discussed in Section 5). Our work also complements current RS, which currently do not factor in consumer inclination for variety seeking at the level of individual product categories for their recommendation strategy.

2   Related Work

2.1   Marketing and IS Motivations for Studying Variety Seeking Behavior

Out of the different possible questionnaire measures we would like to replicate in an information system (ex. such as an RS), we focus in this study on variety seeking behavior. Variety seeking is defined as the "biased behavioral response by some decision making unit to a specific item relative to previous responses within the same behavioral category, due to the utility inherent in variation per se, independent of the instrumental or functional value of the alternatives of items" (Trijp et al. 1996). In

other words, in addition to external circumstances (i.e. external situations like discounts, reviewed in McAlister & Pessemier (1982), or in stock-piling for future preference uncertainty, as reviewed in Kahn (1995)), there is an innate personal characteristic that defines the extent of variety seeking, and that this degree of variety seeking is related to the product category.

This innate variety seeking has implications on other aspects of consumer behavior that is of relevant to retailers. For example, it was shown that consumers who pursued higher variety tend to increase overall consumption quantity (i.e. variety consumption has an additive rather than a substitutive effect) (Kahn & Wansink 2004; Simonson 1990; Read et al. 1995). Variety seekers also tend to be open to promotions (Ailawadi et al. 2001). As such, variety seekers form a potentially valuable group for individual segmentation and marketing initiatives delivered through smartphones. Identifying variety seekers would also be valuable to recommender systems, and complement existing paradigms in that field. Recommender systems (RS) are information systems that identify consumers' product interests and preferences, either explicitly or implicitly, and make recommendations accordingly (Xiao & Benbasat 2007). Many RS attempt to offer consumers products which they have not yet tried (the "cross-sell" paradigm) (Adomavicius & Tuzhilin 2005; Schafer et al. 1999), an appeal to the consumer's desire for variety. RS have focused mainly in the area of e-commerce (rather than physical grocery stores) and generate recommendations based on the content of what consumers have bought or what others have bought (the so called content-based vs. collaborative-based paradigm in RS; for a review see Adomavicius & Tuzhilin (2005) and Xiao & Benbasat (2007)). Content-based systems rely on the consumers' purchase histories and the characteristics (represented as meta-data) of the bought products to infer the right recommendation (for example, recommending products with similar characteristics to ones bought frequently), while collaborative-based systems recommend products based on the purchasing histories of other customers who bought similar products (for example, recommending a product that others buy frequently and is similar to what you buy). Current research in content-based RS focuses on consumer profile building and expanding the set of product characteristics which RS algorithms can infer consumer preferences (Lops et al. 2011), to include also user-generated (i.e. collaborative) content like social tags (Shepitsen et al. 2008; Milicevic et al. 2010); These are dubbed hybrid methods since they are both collaborative and content-based (Burke 2002). However, while these methods try to estimate what *products* a consumer might like, as a methodological and research gap, they ignore the dispersion of the consumer's purchases and hence the *consumer's* variety seeking *inclination,* and also ignore *for which* product *categories* this would occur. Our approach addresses this and thus focuses on the consumer profile building phase of RS, and complements content-based systems by adding the valuable dimension of variety seeking.

2.2 Questionnaire measures in Marketing and IS

The tradition of psychometric questionnaire measures in physical retailing marketing research is well established; frameworks and measures exist for different dimensions of consumers' shopping orientation, such as their price-quality orientation (Garretson et al. 1998; Fisher et al. 2002; Kerin et al. 1992), their price sensitivity (Wakefield & Inman 2003; Dodds et al. 1991), their openness to useful vs. hedonic goods (Voss et al. 2003) and their extent of variety seeking (Trijp et al. 1996; Baumgartner & Steenkamp 1996). The majority of these frameworks and measures have been empirically verified against purchase intentions and in some cases, real purchase data (Baumgartner & Steenkamp 1996) to evaluate their predictive validity. Similarly, recent research in information systems (i.e. mobile shopping aids) have incorporated the influence of consumer beliefs on behavior (van der Heijden 2006; Lee & Benbasat 2010; Kamis et al. 2008), as measured by questionnaire-based constructs.

In spite of the ubiquity of questionnaires in both research and practice, the implementation problems are recognized, as evident by the body of research dealing with non-participation: these range from survey length reduction (Childers & O. C. Ferrell 1979; Bergkvist & Rossiter 2007) to methods of analyzing unanswered questions (Porter 2004; Bosnjak et al. 2005). As noted by Nulty (2008),

participation rate is roughly 30% for online surveys. As such, from an information systems point of view, there is a motivation for having a data-driven method of replicating select insights of the questionnaire measures.

2.3   The Unit of Analysis for Variety Seeking

In terms of the unit of analysis for variety seeking, until recently, past marketing studies have focused on variety seeking as a characteristic trait, independent of product categories (reviewed in Steenkamp & Baumgartner (1992)). The study by Trijp, Hoyer & J. Inman (1996) was an early attempt in developing a psychometric scale that defined the consumer's extent of variety seeking with respect to a particular product category; however, as we will see in Section 4, since a typical grocery retailer can have several hundred distinct product categories, it is not yet possible to practically administer this questionnaire for every single product category for every single consumer. Nor has it been practically possible in the past to be able to take individual action at the product category level with this knowledge. Indeed, past empirical studies of variety seeking have limited their focus on arbitrary and often merely convenient samples of product purchase data available to the researchers; as such their insights on individual levels of variety are arguably not generalizable across all categories. For example, the study by Kahn et al. (1986) examined sandwich bags, wraps, margarine, cereals and soft drinks while the study by Trijp, Hoyer & J. Inman (1996) focused on beer, coffee, hand rolled tobacco and cigarettes, and finally the study by Trivedi (1999) focused on hypothetical purchases of cola. These could hardly be called representative samples of a consumer's product space; indeed, there is an acknowledged research gap in investigating product-category influences on variety seeking (Roehm & Roehm 2004; Tang & Chin 2007; Michaelidou & Dibb 2009). This gap becomes important with smartphones, since we are able to analyze and take marketing action at the granularity of consumers and product categories. In line with this, our approach will characterize individuals' degree of variety seeking for the different product categories in which he shops in.

2.4   Approaches in Enumerating Variety Seeking

In this section we discuss some existing approaches to numerically model variety seeking and their appropriateness for our information system context. Our work is closest in application to the studies from Kumar & Trivedi 2006 and Trivedi (1999), where a variety index was defined for individuals, which described the variation in products a person bought in a product category at a given time. This was then combined with intensity – the number of purchases made in a product category – in order to identify customer typologies.

However, there are two shortcomings with the Trivedi (1999) approach, which we aim to overcome in ours. The first is that the variety index comes from estimating a mathematical model that requires a consumer's stated preference of the different brands; while this could be solicited from a consumer via technology (i.e. a smartphone survey) for some categories, it would be still difficult to scale this for all categories. The second is that the model was estimated in a controlled experimental setting for one product category (soft drinks), which runs counter to our objective of having a lightweight and generalizable description of consumer variety seeking for different categories. The Trivedi model conceptually belongs to what McAlister and Pessemier (1982) notes as the class of models based on attribute satiation. The idea is that by purchasing and consuming a brand, the consumer is "satiated" with that brand's attributes, and therefore seeks out a much different brands. However, this class of model is not easily implemented in practice since consumer perception of attributes is latent and unobservable to the retailer. Accordingly, our approach aims to avoid such requirements in modelling and estimation.

A further class of models for variety seeking are the first order and second order models, which model variety seeking as a changing with time according two competing processes of variety-seeking and inertia seeking, with a state dependency on the most recent purchase (Bawa 1990; Kahn et al. 1986)).

These models differ from earlier zero-order models which assume an individual's variety seeking is state-invariant (Bass et al. 1984; Bass et al. 1976). Although the high order models provide a finer granular description of individual variety seeking, in practice one has to assume a two brand market in order to assure enough degrees of freedom for estimating the model parameters (Bawa 1990; Kahn et al. 1986); as such their model is not appropriate for our multi-brand reality. Furthermore, Bawa (1990) showed that in practice, the descriptive power of the models did not differ much by assuming a higher order; the main advantage of a high order model was added research insight in the underlying mechanisms of variety seeking. As such, our approach will be similar in spirit to the zero-order models.

3 Research Framework

Given the aforementioned conceptual motivation, our work aims to develop a method for estimating an individual's extent of variety seeking in a given product category. We thus need to answer the question: "How many unique products must a customer buy in a product category in a given period of time to be considered a variety seeker?" To address this, we employ two assumptions:

1. Consumers' extent of variety seeking in a given product category can be described with a zero-order model, that is, independent of state (i.e. previous purchases).

2. We assume also that a consumer's pattern of variety seeking in a given product category can be described by counting the number of distinct products that they have bought in a period of time, conditioned against the extent of variety exhibited by other consumers.

In the first assumption, it was noted by Bawa (1990) that in practice, the descriptive power of the models did not differ much by assuming a higher order (i.e. state dependence); the main advantage of a high order model was added research insight in the underlying mechanisms of variety seeking. Hence, we assume a zero order model. In the second assumption, we propose conditioning the variety seeking on the product category, the time period and the behavior of others in the population because in the domain of grocery retailing, it is known that the inter-purchase time (and hence, frequency and variety) of product purchases varies by product category (Rhee & Bell 2002; Leszczyc et al. 2004). For example, a product like milk may have only one to two substitutes and on average a consumer may stick to one variant; as such, a person buying three varieties of milk would be considered variety seeking, whereas for fruit yoghurt, buying three varieties per year might not be considered to be variety seeking, if on average people buy eight distinct flavors per year out of twenty available. Thus, by evaluating the variety of an individual's purchases against the variety shown by the overall population in that category, we are able to control for category-level differences.

3.1 Probabilistic Model Fit for Variety of Goods Purchased

Consequently, the second assumption leads naturally into stochastic counting-based models (Andersen et al. 1993), employed previously in consumer base analysis to successfully predict customer lifetime (Fader et al. 2005; Fader & Hardie 2009). In the models described by Fader et al. (2005), an individual consumer's frequency of purchases at a retailer are counted and used as an input to estimate how long they are likely to continue buying from that retailer. We will use a similar idea by using the count of distinct products bought in a time period to describe variety seeking. To our knowledge, such models have not been applied to the area of variety seeking.

Since count data is discrete, non-negative and has no upper limit, it follows that our probabilistic distribution for modeling the population's variety seeking should be a discrete and positive skewed distribution. For parsimony, we aim for a univariate probability mass functions (PMFs), since this lends itself easily towards an interpretable, one dimensional variety index. PMFs of this type include Poisson, geometric and negative binomial distribution (NBD).

Conceptually, the Poisson distribution is the probability of a given number of events occurring in a fixed interval of time and/or space if these events occur with a known average rate and independently of the time since the last event; modeled on variety seeking, this implies individuals have differing rates of variety seeking and satiation, subject to some stochastic process, and some probability of consuming a certain amount of variety – in-line with the evidence on satiation (McAlister & Pessemier 1982; Chintagunta 1999). This conceptualization is also in-line with Wood & Neal (2009), which noted that individual consumers tend to be habitual in their purchases (i.e. implying the population's variety seeking is positively skewed). The Poisson distribution is given by:

$$\Pr(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$$
(1)

Where k = 0,1,2… and λ is the Poisson shape parameter, with $\lambda > 0$. The distribution gives the probability a consumer has bought k products, given the behavior of the overall population in that product category.

Other distributions are possible, with similarly plausible storylines: The geometric distribution is the probability distribution of the number X of Bernoulli trials needed to get one "success" event, supported on the set of natural numbers. A Bernoulli trial is a random experiment in which there are exactly two outcomes, "success" or "failure", with the probability of success defined the same way each trial (by a fixed parameter). Applied to variety seeking, the Bernoulli trial would be a check at a given time of whether someone is satiated (the "success" event) and will seek no more variety, and the geometric distribution would tell us, in the whole population, how many "failure" events this would take before satiation. The geometric distribution is given by:

$$\Pr(x = k) = (1 - p)^k p$$
(2)

Where k = 0,1,2… and p is the probability of "failure" (i.e. non-satiation/ desire for continued variety).

The NBD proposes a similar story. It models the number of successes in a sequence of Bernoulli trials before a specified (non-random) number of failures (denoted r) occur. For NBD, if we consider "success" to be **non**-satiation, and "failure" to be satiation, then by letting r = 1, then NBD provides the k number of events before satiation for the population. The negative binomial distribution is given by:

$$\Pr(x = k) = \binom{k+r-1}{k}(1 - p)^r p^k$$
(3)

We will test the fit of these models in our study for each product category. Accordingly, we propose the following research question:

> RQ1: Which discrete probability distribution best captures the population's variety seeking?

3.2 Consumer Variety Seeking Heterogeneity across Product Categories

It was reviewed in Section 2 that variety-seeking is a personal characteristic – as such one would expect the degree of variety seeking for a given person to be the same across product categories (i.e. the null hypothesis would hold true for the variety indices of a person). However, we argue that consumer variety seeking will differ between product categories. It is known that the inter-purchase time of product purchases varies by product category (Rhee & Bell 2002; Leszczyc et al. 2004), and hence we argue, so would variety seeking. Furthermore, the attribute-based theory of variety seeking (reviewed in McAlister & Pessemier (1982)) suggest that since different product categories have different attributes which are "consumed", it follows that some categories are more prone to satiation (and hence variety seeking) than others.

Our study would confirm comprehensively whether the variety seeking varies from category to category. Accordingly our research question is:

> RQ2: *Does the variety-index of a customer differ from category to category?*

3.3 Consumer Typologies

As discussed in Section 2, questionnaires are used by marketers to develop customer typologies for better targeting consumers with offers and products. In the case of questionnaires, the variety seeking typology applies at a general level, irrespective of product category. We had argued that our variety seeking indices of a given person at the category level will be heterogeneous across categories. However, in spite of this heterogeneity, it is possible that there is a high-level similarity between consumers, leading to typologies or groups of similar consumers which marketers can address at a coarser granularity. Thus we propose:

> RQ3: *What are higher level similarities between people in spite of their heterogeneous variety seeking between categories?*

In answering this question, by finding a higher level typology of variety seekers, we would achieve the same result as the traditional questionnaire. Hence we would be able to offer information systems such as mobile RS and marketers different granularities of analyzing their customers.

4 Method and Results

4.1 Dataset for Estimation

To address our research questions, we examined a year of point of sale data from our physical grocery retailer partner. The receipt data comes from a complete set of all transactions from a single store, and consists of over 150,000 unique receipts covering a total of two million transactions, with a total of 19,374 unique products sold that year. A transaction refers to the purchase of an item as it appears on the receipt; each transaction event records the name of the item, a timestamp of when it was bought, the European article number (EAN), a receipt ID, the number of units bought, the price per unit, the loyalty card ID of the household who bought it and the product's category as defined by the retailer. Although the customer's age distribution and gender are not known, it is the full dataset of one particular store and was deemed "typical" by our retail partner. Therefore, from this data we are able to construct the purchasing histories of the households, and identify how many items, unique or in total, they bought per category and at what time.

We acknowledge that loyalty card holders might be different from the larger population of all customers and that a given loyalty card might represent a household with different people with different tastes; furthermore it is well acknowledged that consumers often split their purchases between multiple retailers (Rhee & Bell 2002; Leszczyc et al. 2000); these external purchases are unobservable to the retailer. However, since retailers only have measurable data about their own loyalty card holders, our study is relevant and in-line with a retailer's practical limitations. Given the unobservable purchases, in order to determine the "natural" (from the point of view of the consumer) amount of variety and frequency of purchases for different product categories, we estimate our models on consumers who have shopped more than 51 times in our store (N=852). This is based on the observation that across all stores, consumers shop weekly (Rhee & Bell 2002; Leszczyc et al. 2000). We furthermore focus our analysis on products which were available all year around (defined as having been sold at least once per week in the entire store). Since many products are only available few days a year, if we treated these products as "available all year around", this would inflate the apparent number of products available for a consumer (and hence variety) to choose from. We thus analyse 350,000 transaction events. The minimum purchase and product availability filtering are in

line with past models estimated on consumer scanner panel data (Bawa 1990; Gupta 1991; Guadagni & Little 1983; Bucklin et al. 1998).

4.2 Product Categories Definitions

The product categories of the retailer are defined at four hierarchical levels. At the highest (first) level, there are ten categories: "meat and sausages", "fruits and vegetables", "fresh goods", "ingredients", "preserved food", "drinks", "baked goods and sweets", "washing material", "perfume", and "non-food items". These are then broken down into subcategories at the second level – for example, "drinks" would be broken down into the subcategories of beer, coffee, tea, etc., which in turn can be broken down into a third level category ("bottled beer", "canned beer") to finally the fourth and lowest level (different variants of beer such as "draught beer", "wheat beer", etc.). The number of categories at each level is as follows (in brackets we have the number of categories which are sold all year around): there are 66 (50) second level categories, 277 (146) third level categories and 900 (308) fourth level categories.

Although each retailer could have their own arbitrary classification scheme, we note that these categories are mappable to reference categories defined by the widely adopted and open Global Product Classification (GPC) international standard[1]. As such, our results are applicable and can be "translated" across all grocery retailers.

We acknowledge that even the GPC standard categories can seem arbitrary to consumers and might not necessarily match the mental model of the consumer in terms of defining substitutes; hence, the amount of apparent variety seeking can be inflated due to the way the categories themselves are defined, (i.e. a category is too broad). We note however that our model accounts for this because the very definition of "variety" is evaluated with respect to the rest of the population-wide behaviour in that category. We thus test our model fit at the different four levels of product categories to assure its generalizability.

4.3 Probabilistic Model Fit for Variety of Goods Purchased

In order to answer RQ1, we tested the fit of the three proposed models (Poisson, geometric and NBD) to our POS data. Given a product category that is sold all year round, we counted the total number of unique products bought by each consumer in our one-year data set – we denote this vector of frequencies as X. We then fitted a Poisson, geometric and negative binomial distribution to X using maximum likelihood estimation. We computed for each distribution fit the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) in order to compare the goodness of fit of the models. We found that both AIC and BIC gave nearly identical conclusions, so for parsimony the following results show conclusions derived from the AIC. The AIC is also well established for comparing goodness of fit models in various disciplines (Bozdogan 1987; Kamakura & Russell 1989; Andrews & Currim 2003). It measures of how well the model fits (relative to other models) – but there is a logarithmic penalty on the complexity of the model. In other words, AIC balances parsimony with goodness of fit, which is appropriate to our goal of having a parsimonious model.

We fitted the models using first the lowest level product categories, and then repeated the process for the higher level. For each level, there were also categories for which all consumers only bought one item; these categories were not considered in the model fit, since no distribution could be meaningfully fitted. The results are summarized in Table 1, which show that best fitting distribution, irrespective of product category level used, was the Poisson distribution. We therefore adopt the

---

[1] Defined in http://www.gs1.org/gdsn/gpc/what

Poisson distribution for all subsequent analysis. Notably, the geometric function did not emerge once as a best fitting function. RQ1 is addressed.

Category level used	% Categories fitted best to:		# of Categories where no fit was possible (all consumers bought 1 variety of item)
	Poisson	**NBD**	
4th (223 categories fitted)	97%	3%	65
3rd (129 categories fitted)	86%	14%	17
2nd (47 categories fitted)	68%	32%	3

Table 1.   Summary statistics of best fitting distributions using different category levels.

The next step was to determine which category level to choose for subsequent analysis. We note from Table 1 that the lowest level (4th level) category had a high (28%) number of categories for which consumers bought only one variant of a product; we checked then if each of these categories had only one product available. They did. This suggests the 4th level category is too fine and overfits a category 1:1 to a single product. At the same time, we find that the 2nd level is too coarse; for example, the "milk products" category encompasses yoghurt, pudding and butter, which are arguably not substitutes for each other. We therefore choose the 3rd level category (146 categories).

In order to give a more descriptive snapshot and reference point of what the data looks like at a category level, we present in Table 2 a sample of the 146 categories, showing the top 5 product categories in terms of sales volume and the top 5 categories in terms of absolute variety. Table 2 also shows the Poisson parameter, $\lambda_{Poisson}$, which describes the population mean variety seeking. We observe also that these categories with high variety seeking also have a large assortment of available products.

Product Category	# Unique Products	$\mu_{Variety}$ (σ)	$\mu_{Units\ Bought}$ (σ)	Top 5 by variety	Top 5 by sales volume
Soft Drinks	48	3.8 (2.9)	12.0 (16.6)	yes	no
Yoghurt	78	7.2 (5.2)	28.8 (35.1)	yes	yes
Cheese	54	6.1 (4.3)	16.5 (17.6)	yes	yes
Baked Goods	35	4.7 (3.1)	17.7 (21.7)	yes	yes
Milk	11	2.6 (1.4)	33.0 (30.9)	no	yes
Cream	14	3.6 (2.0)	22.3 (22.3)	yes	yes

Table 2.   Sample of variety seeking categories. Note that $\lambda_{Poisson} = \mu_{variety}$

4.4   Consumer Variety Seeking Heterogeneity across Product Categories

4.4.1   Computing the Consumer-Product Variety Index

In order to answer RQ2 and check whether consumer variety seeking is heterogeneous across product categories, we first need to compute the variety index for each person for each product category. With $\lambda_{Poisson}$ computed for each product category, the Poisson distribution (Eq. 1, Section 3.1) would give us the probability a consumer has bought k unique products in that category. The higher the k, the more improbable it is. However, this probability would not be a direct surrogate of variety seeking: ideally, an individual variety seeking index should be monotonically increasing and would describe the person's relative propensity to seek variety compared to the population. The Poisson cumulative distribution function (CDF) fits this and describes the probability someone bought less than or equal to k products. It is given by:

$$\Pr(k \mid \lambda) = e^{-\lambda} \sum_{i=0}^{floor\ k} \frac{\lambda^i}{i!} \qquad (4)$$

Thus someone who has bought a large variety of products relative to the population will also have a high probability, given by (4), that the rest of the population has bought a lesser number of products. As such, the CDF captures the relative degree of variety seeking at the individual level. Using (4), we computed this variety index for all consumers for all product categories.

By using this variety index instead of a raw value of variety of goods purchased, we are therefore able to compare a consumer's variety seeking across product categories, and compare consumers within the same product category.

4.4.2 Confirming Heterogeneity of Variety Seeking Across Product Categories

Having computed the variety indices of a consumer, it follows that if variety seeking was a trait that was category independent, then a given person should have the same extent of variety relative to the rest of the population for each product category (i.e. their variety index should not be significantly different across categories). Furthermore, this would also imply that the population distribution of variety indices for each category would be similar between categories.

The classical test for this statement would be a repeated measures ANOVA, since a consumer is subjected to multiple "treatments" (i.e. shopping in product categories) as stimuli, and the resultant outcomes (i.e. the variety indices) are compared between the treatment groups (the product categories). However, as shown in Section 4.3, since we do not have normally distributed variety indices within groups (the product categories), the assumptions of ANOVA are violated. As discussed in Field et al. (2012), a non-parametric test which can be used in this instance is the Friedman test (Friedman 1937); however a requirement of the Friedman test is completeness of data. In our data set, since not every consumer purchased from every product category, there will be consumer-product category pairs for which no variety index exists. We therefore do the analysis with the Skillings-Mack test (Skillings & Mack 1981), which generalizes the Friedman test to allow for missing data. We found the Skillings-Mack statistic to be significant, $T = 10205$, $p < 0.05$; the product category did significantly change a person's degree of variety seeking. RQ2 is addressed. This therefore justifies the need of computing a variety index at the granularity of person's interaction with a specific category.

4.5 Consumer Typologies

While the findings of RQ2 prove that an overall consumer typology is not sufficient to explain the heterogeneity of variety indices on a category level, we now look at the other way around. In order to answer RQ3, we show that we can use a combination of a consumer's variety indices per category to establish an overall aggregated consumer typology.

Since the variety index of a single category cannot be used to predict the index of another category, we calculate the mean variety index of each customer, based on their category level variety indices. For ease of illustration in this paper, as an example we focus on a subset of the POS data: In particular, we select the top five categories in terms of purchase frequency and only select those customers (N = 547) that have purchased products in all five categories. The selected five categories account for 22% of all purchases and 11% of the revenue – a large and managerially relevant set of categories. We calculate the aggregated mean variety index of the selected five categories per person and visualize the result as histogram on Figure 1(a).
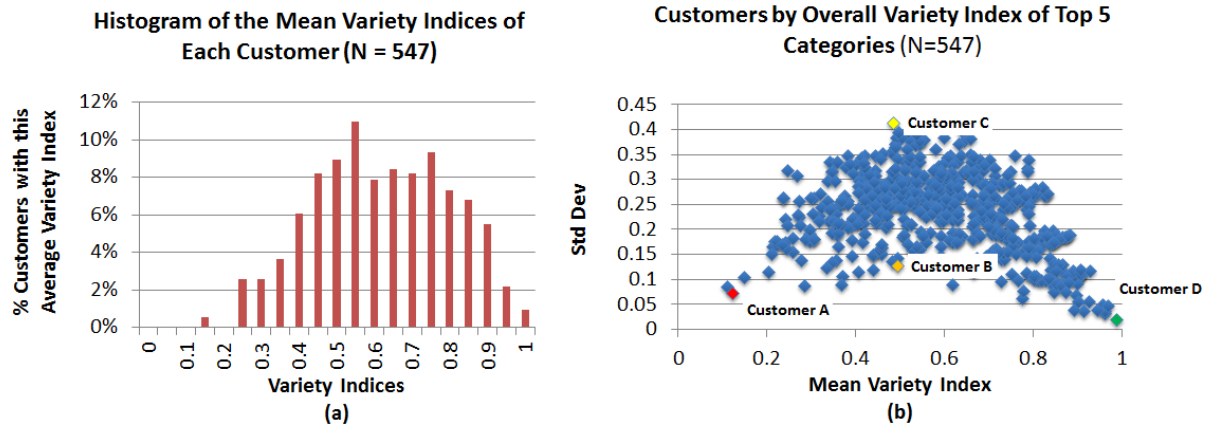
Figure 1. (a) A histogram of the customer's average variety index. (b) A plot of the customer's mean and standard deviations of their variety indices.

In the $\mu_{variety}$ distribution of Figure 1(a), the mean of $\mu_{variety}$ (the grand mean) was 0.59, and the standard deviation 0.18. We characterize those one standard deviation or more below the grand mean to be overall low variety seekers, those one standard deviation or more above the grand mean as overall high variety seekers, and those in between as medium variety seekers. As a result, we have 347 (63.4%) medium variety seekers, 95 (17.4%) low variety seekers and 105 (19.2%) high variety seekers. These proportions are consistent with questionnaire typology data (Trijp et al. 1996), and show that there are some higher level similarities between consumers that can be leveraged for traditional, coarse-granular marketing methods. We also tested whether the distribution of $\mu_{variety}$ can be generalized with a normal distribution. We conducted the Kolmogorov-Smirnov test of normality, which found that the distribution of $\mu_{variety}$, $D(547) = 0.048$, $p < 0.05$ was significantly different from a normal distribution, so the result is not generalizable with a parametric normal distribution. The distribution was also platykurtic (-0.744) and wider than the normal distribution.

To explain the effect of the heterogeneity, we then visualize the population as a scatter plot in Figure 1(b). Each dot corresponds to an individual customer. The x-axis shows the computed mean of the aggregated variety index for each customer, and the y-axis shows the corresponding standard deviation. The result is once again consistent with the findings of RQ2 and shows especially that there are some significant differences between customers and also within customers, especially in the group of "medium" variety seekers. This means that some average variety seekers are actually high variety seekers in some categories and low variety seekers in others, which has significant implication on the success of marketing measures. Meanwhile, we see that customers who exhibit high or low variety seeking are also more consistent in their behaviour, shown by their lower standard deviations. To further illustrate the point, we pick four different customers (shown as coloured dots in Figure 1(b)) and elaborate on their differences in terms of their typology.

	Overall Variety Index &Typology	Variety Index and Typology In:				
		Product Category 1	Product Category 2	Product Category 3	Product Category 4	Product Category 5
Customer A	0.12 (L)	0.11 (L)	0.05 (L)	0.26 (L)	0.07 (L)	0.12 (L)
Customer B	0.49(M)	0.49 (M)	0.68 (M)	0.51 (M)	0.28 (L)	0.51 (M)
Customer C	0.48(M)	0.99 (H)	0.98 (H)	0.26 (L)	0.07 (L)	0.12 (L)
Customer D	0.99 (H)	0.99 (H)	1 (H)	0.95 (H)	1 (H)	1 (H)

Figure 2. Four examples of customers with their overall variety seeking typology and their variety indices in the top five categories. For convenience, variety seeking behavior is marked as L=low, M=medium, and H=high depending on the index value

As Figure 2 shows, customer A and D are fairly consistent in their behavior. In all five categories they exhibit the same level of variety seeking. However, customer B and C are completely different. While their overall typology both amount to medium variety seeking, they do not have a single category on which they show the same behavior. Customer B is a medium variety seeker in four of the five categories whereas customer C is a high variety seeker in two categories and a low variety seeker in three categories.

5 Discussion

Our findings confirm the need for a higher resolution perspective on variety seeking behavior. We prove that the currently dominant approach of using a per-person typology is not sufficient to explain the differences in variety seeking behavior per category. These differences are, however, critical for successful marketing measures in mobile or social commerce. Our presented approach allows for measuring variety seeking behavior on a category level and comparisons between consumers and across categories. We showed that a variety index can be modeled nicely with a Poisson distribution and we provided both confirmation and some characteristics of reference categories based on real point-of-sale data. We acknowledge that often a high-level typology of a consumer is already sufficient for decision making and therefore also provided an method for reconciling the highly granular, consumer-product category perspective with the overall typology used in research and practice today.

5.1 Applicability at Different Granularity Levels of Categories

We confirmed that the variety of goods purchased in a product category for a customer fits a Poisson distribution. This relationship holds true even when we conducted the analysis across different granularity levels of product categories. As such, our relationship is robust against the arbitrariness of product categories as set by the retailer and therefore can be generally applied in different grocery retail settings. The Poisson "storyline" is also intuitive; individuals have differing rates of variety seeking and satiation, subject to some stochastic process, and some probability of consuming a certain amount of variety - in-line with the evidence on satiation (McAlister & Pessemier 1982; Chintagunta 1999). We also saw that the categories where variety seeking was high also had a large assortment of goods available, which confirms past psychological studies that merely having a wide assortment of goods can drive variety seeking (Diehl & Poynor 2010; Kahn & Wansink 2004; Read et al. 1995; Simonson 1990). The product categories which showed high variety of goods sold (ex. soft drinks and fruit yoghurt) also matched the categories selected for study in the literature (Trivedi 1999)

5.2 From Category Level Variety Seeking to Overall Variety Seeking

We finally showed that the category variety indices can also be generalized to determine a person's overall inclination to seek variety, thus replicating the insights of a psychometric questionnaire. Where our approach differs from the questionnaire is that in addition to the overall picture, we are able to quantify individual variety seeking at the category level, thus characterizing the heterogeneity we saw in the spread of standard deviations of $\mu_{variety}$. This means that we expect that our model will be able to better explain consumer receptiveness to try a variety of goods, and also for designing category-level recommendations that cater to a consumers' extent of variety seeking in a particular category.

5.3 Accuracy of results

Although the results are already very promising and the effects are clearly visible, we are convinced that the results can be further improved. In addition to studying the effects in different stores and industries, we consider three factors that are relevant for future research.

In particular, we came across four factors that future studies might include. First, we base our customer identification on the loyalty card number. We acknowledge that this card might be shared within a whole family and thus increase the noise level of our findings. A future study might control for this noise by attempting to identify the demographics behind the loyalty card holder. Second, we assumed that variety seeking behavior remains constant over time on a category level. We took the full data set for twelve months and did not allow for changes over time yet. A future study can try different estimation periods. Thirdly, we excluded seasonally offered products for the analysis and focused only on the top five categories for RQ3. We acknowledge that it might also be of interest to consider seasonal products and all product categories when studying variety seeking behavior, and that extending the automated measuring of variety-seeking to the whole population would be for future work. Finally, in our overall variety seeking aggregation, we used a simple (unweighted) aggregation of variety indices; since consumers have a finite budget, a consumer can only explore a finite number of categories at depth, and as such, the overall variety index could appear much lower due to the presence of many sparsely explored categories. A future study could propose and evaluate a weighted aggregation scheme, where for example categories with low relative spending are weighted less.

6 Conclusion and Future Work

In this paper, we presented one of the first studies to develop a model for measuring variety seeking behavior on a high level of granularity. In particular, we showed how to obtain a high resolution classification based on individual consumers' purchases in individual product categories that is more powerful than established psychological per person typologies. Our study has the following main contributions. Firstly, we contribute to the customer segmentation research stream by providing a novel way for identifying variety-seeking customers and their specific variety-seeking behavior on a category level. Second, for the most important retail categories we characterize the extent of variety seeking and provide a data-driven approach that is easy to operationalize by practitioners – especially for deploying large-scale personalized marketing measures in social or mobile commerce in physical stores. For practitioners, our model is parsimonious and can be estimated on observable consumer purchasing data logged at the PoS; it does not depend on demographics, which are not always available to every retailer due to privacy regulations, nor does it depend on unobservable motivations. Thirdly, we provide a method to reconcile the granular category-level results with per person typologies used today, and thus provide also a starting point for further research. Finally, our work addresses a gap in RS research, which until now was largely based on based on the content of what consumers have bought or what others have bought; our model of characterizing consumers' extent of variety seeking adds another dimension, that of personal inclination to try new goods. This opens up new research opportunities in combining the classical approaches in RS with the insights of this paper. For example, our method could help RS identify *which consumers* seek more variety than others for specific product categories; in turn, once both have been identified, a classical content-based RS can be applied to decide *what product* to recommend.

As a next step, we aim to increase the granularity of our model even further by considering the time domain in addition to the category-level. This would relax the assumption that variety seeking behavior remains constant over time and help to deepen the understanding of the effect of changes in the consumer's behavior. Moreover, we aim to apply the model as a content-based approach for RS in real-time decision making for a smart phone application. Finally, for computing overall variety seeking, a study with a weighted aggregation scheme should be investigated.

References

Adomavicius, G. & Tuzhilin, A., 2005. Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*, 17(6), pp.734–749.

Ailawadi, K., Neslin, S. & Gedenk, K., 2001. Pursuing the value-conscious consumer: store brands versus national brand promotions. *The Journal of Marketing*, 65(1), pp.71–89.

Andersen, P.K. et al., 1993. *Statistical models based on counting processes* 1st ed., Springer.

Bass, F. et al., 1984. An investigation into the order of the brand choice process. *Marketing Science*, 3(4), pp.267–287.

Bass, F., Jeuland, A. & Wright, G., 1976. Equilibrium stochastic choice and market penetration theories: derivations and comparisons. *Management Science*.

Baumgartner, H. & Steenkamp, J., 1996. Exploratory consumer buying behavior: Conceptualization and measurement. *International Journal of Research in Marketing*, 13, pp.121–137.

Bawa, K., 1990. Modeling inertia and variety seeking tendencies in brand choice behavior. *Marketing Science*, 9(3), pp.263–278.

Bergkvist, L. & Rossiter, J.R., 2007. The Predictive Validity of Multiple-Item Versus Single-Item Measures of the Same Constructs. *Journal of Marketing Research*, 44(2), pp.175–184.

Bosnjak, M., Tuten, T.L. & Wittmann, W.W., 2005. Unit (non)response in Web-based access panel surveys: An extended planned-behavior approach. *Psychology and Marketing*, 22(6), pp.489–505.

Bucklin, R., Gupta, S. & Siddarth, S., 1998. Modelling the effect of purchase quantity on consumer choice of product assortment. *Journal of Forecasting*, 301, pp.281–301.

Burke, R., 2002. Hybrid recommender systems: Survey and experiments. *User modeling and user-adapted interaction*, 12, pp.331–370.

Chandon, P., Wansink, B. & Laurent, G., 2000. A Benefit Congruency Framework of Sales Promotion Effectiveness. *Journal of Marketing*, 64(4), pp.65–81.

Childers, T.L. & O. C. Ferrell, 1979. Response rates and perceived questionnaire length in mail surveys. *Journal of Marketing Research*, 16(3), pp.429–431.

Chintagunta, P., 1999. Variety seeking, purchase timing, and the "lightning bolt" brand choice model. *Management Science*, (November 2013).

Diehl, K. & Poynor, C., 2010. Great Expectations?! Assortment Size, Expectations, and Satisfaction. *Journal of Marketing Research*, 47(2), pp.312–322.

Dodds, W., Monroe, K. & Grewal, D., 1991. Effects of price, brand, and store information on buyers' product evaluations. *Journal of Marketing Research*, 28(3), pp.307–319.

Fader, P.S. & Hardie, B.G.S., 2009. Probability Models for Customer-Base Analysis. *Journal of Interactive Marketing*, 23(1), pp.61–69.

Fader, P.S., Hardie, B.G.S. & Lee, K.L., 2005. Counting Your Customers? the Easy Way: An Alternative to the Pareto/NBD Model. *Marketing Science*, 24(2), pp.275–284.

Field, A., Miles, J. & Field, Z., 2012. *Discovering Statistics Using R* 1st ed., SAGE Publications Ltd.

Fisher, D., Burton, S. & Garretson, J.A., 2002. Antecedents of private label attitude and national brand promotion attitude: similarities and differences. *Journal of Retailing*, 78(2), pp.91–99.

Friedman, M., 1937. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association*, 32(200), pp.675–701.

Garretson, J.A. et al., 1998. A Scale for Measuring Attitude toward Private Label Products and an Examination of its Psychological and Behavioral Correlates. *Journal of the Academy of Marketing Science*, 26(4), pp.293–306.

Guadagni, P.M. & Little, J.D.C., 1983. A Logit Model of Brand Choice Calibrated on Scanner Data. *Marketing Science*, 2(3), pp.203–238.

Gupta, S., 1991. Stochastic models of interpurchase time with time-dependent covariates. *Journal of Marketing Research*, 4, pp.0–15.

Van der Heijden, H., 2006. Mobile decision support for in-store purchase decisions. *Decision Support Systems*, 42(2), pp.656–663.

Kahn, B., 1995. Consumer variety-seeking among goods and services: An integrative review. *Journal of Retailing and Consumer Services*, 2(3), pp.139–148.

Kahn, B., Kalwani, M. & Morrison, D., 1986. Measuring variety-seeking and reinforcement behaviors using panel data. *Journal of Marketing Research*, 23(2), pp.89–100.

Kahn, B. & Wansink, B., 2004. The influence of assortment structure on perceived variety and consumption quantities. *Journal of Consumer Research*, 30(4), pp.519–533.

Kamis, A., Koufaris, M. & Stern, T., 2008. Using an attribute-based decision support system for user-customized products online: an experimental investigation. *MIs Quarterly*, 32(1), pp.159–177.

Kerin, R.A., Jain, A. & Howard, D.J., 1992. Store Shopping Experience and Consumer Price-Quality-Value Perceptions. *Journal of Retailing*, 68(4), pp.376–397.

Kumar, A. & Trivedi, M., 2006. Estimation of variety seeking for segmentation and targeting: An empirical analysis. *Jorunal of Targeting, Measurement and Analysis for Marketing*, 15, pp.21–29.

Lee, Y.E. & Benbasat, I., 2010. Interaction design for mobile product recommendation agents. *ACM Transactions on Computer-Human Interaction*, 17(4), pp.1–32.

Leszczyc, P.P., Sinha, A. & Sahgal, A., 2004. The effect of multi-purpose shopping on pricing and location strategy for grocery stores. *Journal of Retailing*, 80(2), pp.85–99.

Leszczyc, P.P., Sinha, A. & Timmermans, H.J.., 2000. Consumer store choice dynamics: an analysis of the competitive market structure for grocery stores. *Journal of Retailing*, 76(3), pp.323–345.

Lops, P., de Gemmis, M. & Semeraro, G., 2011. Content-based recommender systems: State of the art and trends. In *Recommender Systems Handbook*. Springer, pp. 73–105.

McAlister, L. & Pessemier, E., 1982. Variety seeking behavior: An interdisciplinary review. *Journal of Consumer research*, 9(3), pp.311–322.

Michaelidou, N. & Dibb, S., 2009. Brand switching in clothing: the role of variety-seeking drive and product category-level characteristics. *International Journal of Consumer Studies*, 33(3), pp.322–326.

Milicevic, A.K., Nanopoulos, A. & Ivanovic, M., 2010. Social tagging in recommender systems: a survey of the state-of-the-art and possible extensions. *Artificial Intelligence Review*, 33(3), pp.187–209.

Nulty, D.D., 2008. The adequacy of response rates to online and paper surveys: what can be done? *Assessment & Evaluation in Higher Education*, 33(3), pp.301–314.

Porter, S.R., 2004. Raising response rates: What works? *New Directions for Institutional Research*, 2004(121), pp.5–21.

Read, D. et al., 1995. Diversification Bias : Explaining the Discrepancy in Variety Seeking Between Combined and Separated Choices. *Journal of Experimental Psychology: Applied*, I(1), pp.34–49.

Rhee, H. & Bell, D.R., 2002. The inter-store mobility of supermarket shoppers. *Journal of Retailing*, 78(4), pp.225–237.

Roehm, H.A.J. & Roehm, M.L., 2004. Variety-seeking and time of day: why leader brands hope young adults shop in the afternoon, but follower brands hope for morning. *Marketing Letters*, pp.213–221.

Schafer, J., Konstan, J. & Riedi, J., 1999. Recommender systems in e-commerce. *EC '99 Proceedings of the 1st ACM conference on Electronic commerce*, pp.158–166.

Shepitsen, A. et al., 2008. Personalized recommendation in social tagging systems using hierarchical clustering. *Proceedings of the 2008 ACM conference on Recommender systems - RecSys '08*, p.259.

Simonson, I., 1990. The effect of purchase quantity and timing on variety-seeking behavior. *Journal of Marketing Research*, (May), pp.150–163.

Skillings, J. & Mack, G., 1981. On the use of a Friedman-type statistic in balanced and unbalanced block designs. *Technometrics*, 23(2), pp.171–177.

Steenkamp, J. & Baumgartner, H., 1992. The role of optimum stimulation level in exploratory consumer behavior. *Journal of Consumer Research*, 19(3), pp.434–448.

Tang, E. & Chin, I., 2007. Analyzing variety seeking behavior using panel data. *Journal of International Consumer Marketing*, (November 2013), pp.37–41.

Trijp, H.C.M. Van, Hoyer, W.D. & Inman, J., 1996. Why Switch? Product Category Level Explanations for True Variety-Seeking Behavior. *Journal of Marketing Research*, 33(3), p.281.

Trivedi, M., 1999. Using Variety-Seeking-Based Segmentation to Study Promotional Response. *Journal of the Academy of Marketing Science*, 27(1), pp.37–49.

Urbany, J.E., Dickson, P.R. & Kalapurakal, R., 1996. Price Search in the Retail Grocery Market. *Journal of Marketing*, 60(2), pp.pp. 91–104.

Voss, K.E., Spangenberg, E.R. & Grohmann, B., 2003. Measuring the Hedonic and Utilitarian Dimensions of Consumer Attitude. *Journal of Marketing Research*, 40(3), pp.310–320.

Wakefield, K. & Inman, J., 2003. Situational price sensitivity: the role of consumption occasion, social context and income. *Journal of Retailing*, 79, pp.199–212.

Wood, W. & Neal, D.T., 2009. The habitual consumer. *Journal of Consumer Psychology*, 19(4), pp.579–592.

Xiao, B. & Benbasat, I., 2007. E-commerce product recommendation agents: use, characteristics, and impact. *MIS Quarterly*, 31(1), pp.137–209.