A Multi-category Approach to Modeling Consumer Preference Evolution: The Case of Sporting Goods

Nara Youn
Henry B. Tippie College of Business
University of Iowa
Marketing Department
Iowa City, IA, 52245
nara-youn@uiowa.edu

Inseong Song
Hong Kong University of Science and Technology
Department of Marketing
Clear Water Bay, Kowloon, Hong Kong
mksong@ust.hk

Douglas MacLachlan
Michael G. Foster School of Business
University of Washington
Marketing and International Business
Seattle, WA, 98195-3200
macl@u.washington.edu

June. 2008

Abstract

In this research we examine how consumers' brand preferences and price sensitivities evolve as their levels of experience increase in outdoor sports gear categories. Using a Hidden Markov modeling framework, we investigate how consumers evolve along discrete latent states that represent their experience and knowledge levels, and how this corresponds to changes in preference parameters. In our empirical estimation and evaluation of the model we utilize customer membership data from a large outdoor and sporting goods retailer. Existing single-category Hidden Markov approaches require a rich purchase history and do not work well for the types of durable product categories in an area such as sport climbing, as typical customers make relatively few purchases in any single category, even over long periods of time. For this reason we use a multi-category framework in which unobservable experience levels are inferred by leveraging information from multiple categories. Using a Hierarchical Bayes estimation, we account for initial consumer heterogeneity as customers first take up the sport at different skill levels. This also allows us to account for consumer heterogeneity in transition probabilities as customers accumulate experience and knowledge at different rates. We model the latent states to follow a Markov chain with a transition probability matrix that is specific to individual customers, and is a function of their cumulative purchases in each category. Thus, we are able to identify the typical brand(s) purchased by customers at different stages of their experience. We find empirical evidence that customers who are beginners or are new to sport climbing will prefer certain brands while customers who have more experience or knowledge of the sport will prefer other brands. The multi-category approach allows us to determine which product category is most indicative of a stage of a consumer's evolution.

1. Introduction

Understanding how consumer preference and choice evolve over time through consumer learning is key to the strategic management of customers and products. In fact, the modeling of brand choice dynamics has been a focus of marketing researchers for more than a decade. A significant literature documents a variety of approaches to modeling the evolution dynamics that influence consumer choice processes. This previous research, however, suffers for the scarcity of marketing datasets that have sufficient longitudinal depth to fully showcase the internal dynamics underlying the preference formation and informed decision-making of consumers – i.e., the latency of the process. One approach to capturing the latent structure of these dynamics that has shown promise is state space modeling.[1]

In this paper, we propose using a Hidden Markov Model (HMM)[2] as a dynamic learning, state space model to analyze the evolution of consumer preference in a particular sporting goods category. The Markovian states of our HMM comprise a finite set of hidden experience states for consumers. Consumer experience is modeled structurally through an integrated framework that links the unobserved but evolving latent states of experience with the realized outcome of brand choice. The transitions between experience states for a consumer are determined by the consumer's participation in sports activities that are measured by the cumulative purchase of products required to engage in the sport. The model is calibrated using purchase data over a six-year period, across three categories of sports equipment.

The key research questions that we address in our study are the following: 1) *How does a consumer evolve over time? How can we classify consumers in a managerially meaningful way?* In our research, using latent discrete states, we identify different consumer segments based on their levels of experience in a particular sport. 2) *How do consumers develop their preferences over time? Do they stick to the same brand or switch brands over time? Which brands are preferred by inexperienced consumers and*

---

[1] See Lachaab et al. 2006 for a review of state space approaches to modeling preference dynamics in marketing.

[2] A Hidden Markov Model (HMM) is a model of a stochastic process that can not be observed directly but,can only be viewed through another set of stochastic processes that produce a set of observations (Rabiner 1989). A HMM can be seen as a mixture model whose mixing distribution is a finite state Markov chain (Scott 2002).

which by experienced consumers? Can we identify typical brand(s) at different stages of customer experience and set up a series of communications that follow the sequential patterns? 3) *Do consumers become price insensitive over time? Is the finding in existing research on consumer preference evolution (Heilmann et al. 2000) consistent with what we observe in sporting goods categories?* 4) *In the multi-category case, what is the most representative category that can closely indicate consumer evolution?*

In our study, we identify three latent experience states (beginner, intermediate, and advanced), and estimate the marginal impact of the cumulative purchases made by a customer in separate gear categories as the customer transitions between these states. For each experience state, we observe different choice behavior among customers. Price sensitivities and the influence of the last brand purchased, or inertia, also vary over the different states.

Our proposed model for the case at hand has several advantages over previous attempts to characterize brand choice dynamics. First, the HMM allows us to segment consumers according to their experience and skill levels in their activities – in this instance, sport climbing. This segmentation is helpful in investigating the varying impact of price, brand loyalty, and cumulative purchases of related equipment on the probabilities of subsequent brand choices across groups of consumers. It provides insight into what would be a strategic targeting of consumers at different stages or experience levels. Unlike the more static, traditional segmentation models (Kamakura and Russell 1989), HMM dynamically accommodates consumers who switch membership across segments, and thus allows for changes in consumer behavior over time and the phenomenon of consumer evolution (Brangule-Vlagsma, Pieters, and Wedel 2002). With HMM, consumers may switch from one segment to any other in a systematic way (i.e., through a transition matrix with estimable transition probabilities) over time. Second, HMM allows for potential changes in consumer choice behavior that may specifically resultfrom learning. A customer gains knowledge on equipment brands with increased experience, just as that customer will continue to acquire the physical and mental skills demanded of the sport with each climb. This aspect of continual learning necessarily impacts a consumer's price sensitivity and brand loyalty, and may ultimately lead to changes in choice behavior. HMM provides a simple and concise way to account for the possibility of these changes.

Third, HMM allows for the control of consumer heterogeneity that characterizes the sport climbers, and those who consume sporting good more generally. When attempting a sport for the first time, customers will possess different initial states of fitness, knowledge, and related experience or skills. Moreover, customers learn about the products available and about the sport itself over time at different rates. HMM provides an effective way to account for consumer heterogeneity in initial conditions, as well as in transition probabilities, in order to account for those who begin at various levels and accumulate experience and knowledge at different rates.

Our HMM of preference evolution makes real contributions to the literature related to HMM applications in marketing. First, substantively our application of HMM to the sporting goods industry is unique. In our review, no previous study has documented the dynamics of consumer preference evolution in sporting goods. Second, methodologically we utilize the approach developed by Chib (1996) in which unobserved states are simulated. As well-documented in Scott (2002), there are several reasons to prefer sampling states instead of integrating them out.[3] Third, our HMM uses a non-stationary and non-homogenous Markov chain, which means that transitions between states are determined by a set of time-varying covariates. With the exception of Netzer, Lattin, and Srinivasan (2007), all marketing HMM applications utilize stationary and homogenous transitions. The incorporation of covariates into the Markov chain is worthwhile since the resulting Markov chain has a useful substantive interpretation (MacDonald and Zucchini, 1997). In our case, the transition probability is a function of weighted average of the number of cumulative purchases made by the customer in each gear category. This specification allows us to identify the category which is the most indicative of consumer preference evolution.

1.1 Sporting Goods

---

[3] The first reason is that Metropolis-Hastings algorithms perform poorly when the dimension of a parameter vector is large. Most candidate draws are rejected so the algorithm moves slowly. In this case, sampling the hidden states is preferred to integrating them out. The second reason is that effectively drawing the states may accelerate mixing. Finally, the sampled states can be used for inference and as diagnostics for model adequacy and MCMC convergence.

Today's growing popularity of adventure sports and outdoor recreational activities has led to a boom in the related equipment and outfitting industry. According to a 2006 Outdoor Industry Foundation report, Americans alone are now spending $289 billion for trips and equipment a year[4]. As touted by the National Sporting Goods Association, sales of specialty outdoor merchandise (equipment, apparel and footwear) reached $7.62 billion in 2006, and sales have grown 47% since 2000, representing a 6.6% annualized growth rate[5]. Sport climbing is one activity that has captured the imagination of new adventure enthusiasts, as evidenced by the increasing number of designated climbing sites and facilities, even among downtown health clubs and retail shopping areas. In 2006, an estimated 1,897,000 people in the U.S. participated in traditional mountain rock climbing at least once[6].

In this section, we discuss certain unique characteristics of the sporting goods industry, and sport rock climbing in particular, and provide our research motivation for this study. Table 1 contains a checklist for the gear required for a type of climbing known as top-roping, and a checklist of additional items required for another type known as lead climbing.

[Table 1] Rock Climbing Gear Checklists

Top-Roping	Lead Climbing
rock shoes, harness, helmet, non-locking carabiners (6-8), locking carabiner (1), belay device, rope (9.8mm-11mm), runners (3-6), tubular webbing (20-60 ft.), medium to large stoppers/ hexes/ nuts (10-12), nut extraction tool, rope bag, chalk bag, chalk, leather belay gloves, tape	non-locking carabiners (12-28), large locking carabiners (1), locking carabiners (2-6), quickdraws or quickdraw sets (6-8), runners-single (6-10) and double (2-4), passive camming devices (3-6), wired stoppers-small (6-12), spring loaded camming devices (6-10)

(Source: REI Website www.rei.com "Climb Expert Advice")

4 "The Active Outdoor Recreation Economy" published by the Outdoor Industry Foundation, Fall 2006.
5 "Outdoor Industry Association's US Specialty Outdoor Market Study" National Sporting Goods Association Research Newsletter, August 27, 2007.
6 "The Next Generations of Outdoor Participants," published by the Outdoor Industry Foundation, 2007.

Top-roping is a form of rock climbing where the safety of the climber is provided via a rope that is strung from the top of a cliff and anchored securely. The security of an anchor at the top allows climbers a certain amount of leeway to make mistakes. Top-roping is thus a recommended introduction to climbing. Lead climbing is a more advanced form of climbing, and entails having a climber who leads and sets anchors on the way to the top of a climbing course–a much more difficult and precarious activity.

The gear list for top-roping on the left column of Table 1 includes: rock shoes, harness, helmet, non-locking carabiners, locking carabiner, belay device, rope, runners, tubular webbing, medium to large stoppers/ hexes/nuts, nut extraction tool, rope bag, chalk bag, chalk, leather belay gloves, and tape. Lead climbing requires all the gear of top-roping, but with the specifications and additions indicated in the right column of Table 1. Still necessary are rock shoes, harness, and helmet, but as listed, non-locking and locking carabiners are needed in greater quantities. Passive camming devices, wired stoppers, and spring-loaded camming devices are also listed as requirements.

The lists contain items we would expect to find in a customer’s shopping basket. If the customer is a beginner, relatively new to rock climbing, or plans to do only easy climbing tasks, then we would expect the customer to buy the items listed only in the left column. If the customer is a more experienced climber or plans to do more difficult climbing tasks, we would expect to see some of the items from the right column in the shopping basket. Additionally, we expect the content of the basket to change over time as a climber gains experience and learns more about rock climbing. Rock climbing is heavily gear or hardware dependent, and learning about different climbing hardware and different brands and styles can take years of experience. At the same time, climbers have the highest incentive to learn about gear and keep themselves informed about new products; having reliable gear and the knowledge to use it properly is clearly a matter of importance.

In rock climbing, brand preference within a single product category evolves with the customer’s experience level as well -- for example, rock shoes differ in their shapes and stiffness, and preferences for differences in these qualities will vary with each customer. All-purpose shoes, which are usually recommended for beginners, are comfortable and soft. All-purpose shoes, typically cut high to provide ankle support, are designed to be comfortable as well as protective. So-called performance shoes, which are for the more advanced climbers, are relatively stiff and uncomfortable. They are cut low for added ankle flexibility and lighter weight, and are designed to fit tightly for maximum rock-sensitivity and control. Rock shoe manufacturers usually carry a line of different brands that will vary in shapes and stiffness. Here, we anticipate brand preference to evolve along with the customer's level of knowledge and skill.

There are the other sporting goods product categories where we would expect category experience to have a significant impact on choice probability. Industry representatives indicate that they observe changes in brand preference with increased experience in other sporting goods categories such as cycling, skiing, and kayaking. In kayaking, for example, paddling-trip preferences (day trips, expeditions, family recreation) and the type of water determine what kind of boat a paddler would choose to purchase. A kayak's design and the materials used in its construction affect how it handles in various situations.

The evolution of consumer brand preference has also been observed among other specialized industries – notably, in the sales of baby diapers. Heilman et al. (2000) find that when consumers are new to a market – in this case, first time parents – they tend to choose leading national brands (top dogs) due to their lack of category experience. As parents gain experience over time and become familiar with the category, they typically switch to cheaper, private (generic) label products (underdogs). In rock climbing and other sports, however, customers switch up to more complicated and advanced products over time, while for diapers, customers switch down to cheaper brands over time. We expect to observe similar preference evolution for home improvement tool kits as well. Every homeowner tends to have a supply of basic tools around the house. Consumers would tend to purchase cheaper, all-purpose tool boxes when they begin to use hand tools for minor home repairs. When they gain more experience and grow to enjoy large do-it-yourself projects, they would want more expensive professional tool boxes.

In table wine, perhaps customers start with less expensive ones and over time switch up as they gain confidence in their abilities to judge quality (e.g. – a wine expert's prescriptive advice to the budding wine lover: "Do start with simple and inexpensive wines, and work your way up to the powerhouse bottles. Do try a variety of wines. Trying everything is the only way to build your sensory memory and discover your own tastes. You'll never make any progress with wine if you stick to the same Chardonnay or Cabernet Sauvignon, no matter how much you like them.[7]") In pharmaceutical products, there is uncertainty in the efficacy of a drug. We would expect patients' brand preferences to change to generic labels as they accumulate more knowledge about the drug category. In general, when a product category involves considerable risk perceived by novices (i.e., in terms of physical, economic, or psychological harm), one would expect consumers to be willing to pay a high price initially, but seek lower priced alternatives as they gain knowledge and experience.

[Table 2] Key Characteristics of Sporting Goods Categories

- Consumer Learning
    - Customers' purchase decisions are influenced by the extent of their experience or knowledge of the sport
    - Some products are better suited for beginners and some for more advanced customers

- Different Starting Point
    - Customers' initial fitness and knowledge levels vary when they attempt a sport for the first time

- Different Rate of Learning
    - Customers learn about the products available and about the sport itself over time at different rates

As shown in [Table 2], sporting goods categories have unique characteristics that are not likely to be observed in typical consumer packaged goods categories. First, we expect the purchase decisions of customers to be affected by their degree of experience, knowledge, and skill in a sport. Second, customers vary in their experience, knowledge, and skills when they initially begin playing a sport. When customers play a sport for the

[7] Blue, A. D. 2002, The essential wine guide (accessed August 27, 2007) http://eat.epicurious.com/drink/ewg/index.ssf?/drink/wine/ewg/ewg.html)

first time, some will start with easy tasks, but others will start with a higher level of difficulty because they are physically more fit or have some prior knowledge of the sport (or related sports). It is also possible that some customers will take up both easy and difficult tasks at approximately the same time – for example, a climber may choose to go bouldering (climbing close to the ground without a rope) or top-roping with the family one weekend, but soon after go on a week-long trip for a more difficult, traditional lead climb. Differences in knowledge and skills among consumers at any point may result from their experiences before or after their first taking up the sport. It is also possible that at the point when our data period begins, consumers had already accumulated a degree of experience or knowledge. Since our data is both left- and right-truncated, we are not able to observe when consumers first participate in the sport. Thus, we expect to find customer heterogeneity at the initial starting level. Third, we also expect to observe customer heterogeneity in the rate of their individual progress in the sport. Some customers will advance to levels of greater experience and skill quickly and others will remain at the initial state over a relatively long time. Customers learn over time about the products available and about the sport itself at different rates.

1.2 Modeling Challenges

One of the principal challenges to analyzing the relationship dynamics between learning and behavioral changes is the latency of learning. Latency of learning in prior research does not allow one to explicitly analyze how and why individuals differ in learning from activities, and whether the effect of learning activities on an individual changes over time. In this study, we explicitly conceptualize a learning/experience state which allows us to analyze these issues in greater detail.

In this paper, we utilize a *Hidden Markov* modeling (HMM) framework that relates the unobserved but evolving knowledge, skills, and experience levels of a consumer to the realized outcomes of learning – that is, the observed purchase behavior. The latent state is an aggregate measure of a consumer's knowledge, skills, and experience relevant to the product and the sport. The discrete (latent) experience states are inferred through a consumer's participation in sporting activities that is observed through her frequency of purchases of items required for the sport. Through this model, we investigate the marginal impact of purchases made in multiple product categories that affect the transitions of a customer between these hidden states. The model demonstrates how consumers evolve along discrete latent states that represent their experience and knowledge levels, and how their preference parameters change accordingly over time.

Why, then, do we require a multi-category framework? In the world of sporting goods, typical customers make only a few purchases in a single category, even over a relatively long period. (Climbers will not buy rock shoes as often as they buy orange juice.) Existing single category HMM approaches that require a rich purchase history do not work well for these types of durable product categories. Thus, we use a *multi-category framework* in which unobservable experience levels are inferred by leveraging information from multiple categories.

In light of the unique characteristics of sporting goods categories discussed in section 1.1, we incorporate customer heterogeneity in our initial conditions and also in our transition probabilities in order to account for customers who begin climbing at various levels, and accumulate experience and knowledge at different speeds. As a by-product, we are able to identify typical brand(s) purchased at different stages of customer experience. The multi-category approach also allows us to determine which product category is most indicative of the stage of consumer evolution.

1.3 Related Literature

Publications related to our present research address the topics of brand preference evolution (e.g. Heilman et al. (2000), Hidden Markov models (e.g. Moon, Kamakura, and Ledolter, 2007, Netzer (2007), and multi-category analysis (e.g. Ainslie and Rossi (1998).

As discussed in the introduction section, Heilman et al. (2000) investigate the evolution of brand preference among customers new to a market. They review prior behavioral studies on the impact of category experience and learning on dynamic choice processes. The two cognitive factors identified as drivers of choice in the paper are perceived risk and information search. They propose a three-stage process: (1) upon entering a category, typical consumers will limit their information search to top dog brands since the perceived risks associated with underdog brands are high; (2) as a consumer’s willingness to try an alternative brand increases, so does the probability of choosing an underdog brand increase; (3) as a consumer's experience in the category increases, loyalty to the brand develops that provides the greatest utility. Heilman et al. (2000), however, is a single category model, and does not address consumer heterogeneity in evolution rates. More recently, Lachaab et al. (2006) address preference evolution using a hierarchical Bayesian state space model of discrete choice.[8] Here, temporal variability in preference and cross-sectional heterogeneity are incorporated in a multinomial probit model via a correlated sequence of population distributions, i.e. the multinomial probit parameters are drawn from a population distribution that evolves from one period to another.

Previous research on preference learning has also demonstrated that early selections (as starting points) have a lasting impact on future preferences (Hoeffler, Ariely, and West 2006, Hoeffler and Ariely 1999). Hoeffler, Ariely and West (2006) argue that preferences are shaped by initial experiences, and later preferences are path dependent and describe a biased search process. The mechanism is composed of three elements: starting point, favorableness of early experience, and myopic search. They also show that when compared with an unfavorable initial experience, a favorable initial experience leads to fewer test trials in subsequent choice incidence, leading to a lower level of preference development. This provides a theoretical understanding as to why the last brand purchased has a significant impact on preference formation over time. They also argue that the shaping of preferences is not based on psychophysical adaptations or changes in taste. Rather, it is driven by people's ability to control and shape their future experience. The "last brand purchased" effect will also be related to the uncertainty of a choice situation. When uncertainty is high, there will be less variety seeking, whereas when uncertainty is low, more variety seeking. Campbell and Goodstein (2001) show that people have a greater preference for the norm, such that they become more conservative in their evaluations when goals are associated with higher risk.

MacDonald and Zucchini (1997) provide an overview of Hidden Markov Models and examples of applications in multiple fields, including geophysics, finance, and climatology. HMM has only recently been applied to marketing problems. Using HMM,

[8] HMM is closely related to Bayesian state space model (West and Harrison, 1997) in the way that Bayesian state space model assumes that a sequence of unobserved hidden states follow a Gaussian process (Scott, 2002).

Moon, Kamakura, and Ledolter (2007) manage to capture the competitor's unobserved promotional efforts as a latent variable in their promotion response model and find that the HMM reduces biases in the own- and cross-promotion parameters. In the context of customer relationship management (CRM), Netzer, Lattin, and Srinivasan (2007) utilize HMM to model the effect of latent alumni-university relationship states on gift-giving behavior. They find the predictive validity of HMM superior to that of the latent class model and the recency frequency model commonly used in CRM analysis. Using click-stream data from an online bookseller, Montgomery et al. (2004) analyze web-browsing behavior by categorizing the path, or the sequence of web pages, viewed by users. The paths reflect a user's unobservable goals which can be used to predict future movements at a website. The latent states in HMM captures memory or "path dependence" in the sequences chosen. In Liechty, Pieters, and Wedel (2003), the latent states indicate brain switching between global and local attention and the HMM model is calibrated on eye-movement data from a study of customers' attention to print advertisements. Du and Kamakura (2006) take the stages of the household life cycle as unobservable states and identify the most common types of households and the typical sequences in which households move through the life paths. Brangule-Vlagsma, Pieters, and Wedel (2002) apply HMM to dynamic value segmentation in which the segment specific value-utility parameters are constant over time but the consumers are allowed to switch from any segment to any other in a systematic way. They observe substantial switching among segments over time.

The majority of prior research on multi-category analysis has focused primarily on investigating commonalities in marketing mix sensitivities. Ainslie and Rossi (1998) find substantial correlations in household price, display, and feature sensitivity.[9] Erdem (1998) finds consumers' preferences for a brand name are correlated across categories. In our model, categories share customers' levels of category knowledge and we are interested in studying how customers' market baskets change over time as they gain experience.

---

[9] Ainslie and Rossi (1998) modeled the similarity in household behaviors across categories. If we have zero order (instead of the first-order Markov chain), then it would be a discrete heterogeneity version of Ainslie and Rossi.

The paper is organized as follows. We describe the data and the industry background in section 2. Section 3 describes the model formulation. We present our results in section 4 and conclude the paper in section 5.

2. Data

For the purposes of empirical estimation and evaluation of our proposed model, we use data from a large outdoors and sporting goods retailer. This particular company is a dominant player in the market with a near local monopoly. The company is also a consumer cooperative in that the company returns a portion of its profits to its member-customers in the form of a patronage dividend (members receive rebates amounting to 10% of their yearly purchases). Member-customers express a high level of satisfaction with the retailer, with loyalists representing 85% of the company's total sales.

Product categories in the overall dataset include: Camping/Hiking, Rock Climbing, Cycling, Paddling, Snow Sports – Skiing/Snowboarding, and Apparel. Though we believe customer experience plays a significant role in the evolution of brand preference in most, if not all, of these categories, we focus our attention on rock climbing for the compact size and relative simplicity of the dataset associated with this specific sporting activity. For our model calibration, we randomly chose 350 customers from a pool of 528 total customers for whom complete purchase histories were available over a period running from January 1998 to July 2005.

Within the product category of rock climbing gear, we concentrate on three subcategories of equipment: carabiners, passive and active protection, and harnesses. Carabiners are metal snap-links that are used to hitch together ropes and other objects, whether they be climbers to their ropes, ropes to a piece of protection equipment, or a collection of protection pieces to a climbing harness. They are constructed of strong, lightweight materials so that climbers can carry a supply of them without being weighted down, and come in a variety of sizes and styles. Protection includes the bolts and anchors that are set into the climbing surfaces or rock face, and physically link the climber's rope to the rock. A harness is something a climber wears to link the climber's body to the rope.

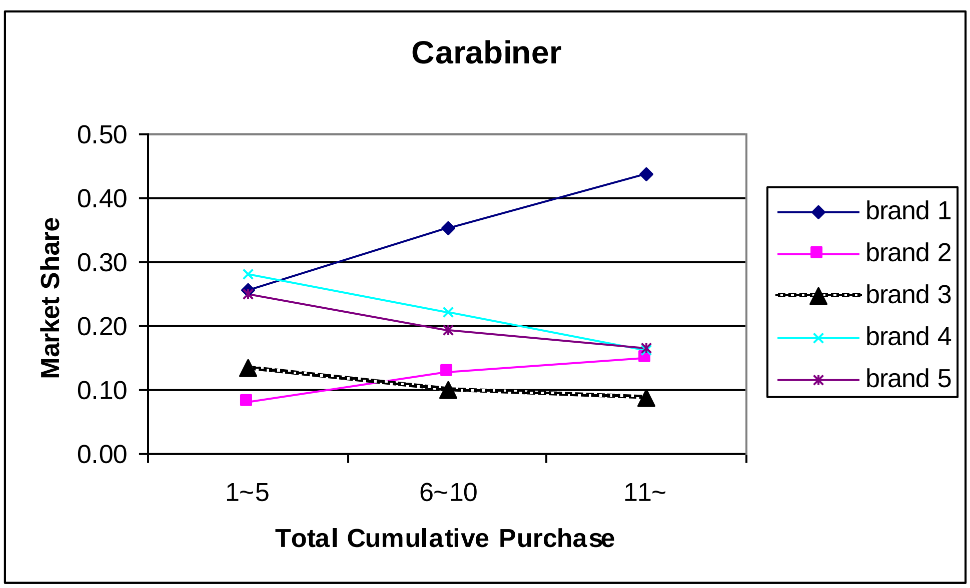
[Table 3] lists the number of observations in the three product subcategories for the customers in our dataset, as well as the market share of five major brands in each subcategory.

[Table 3] Data Description

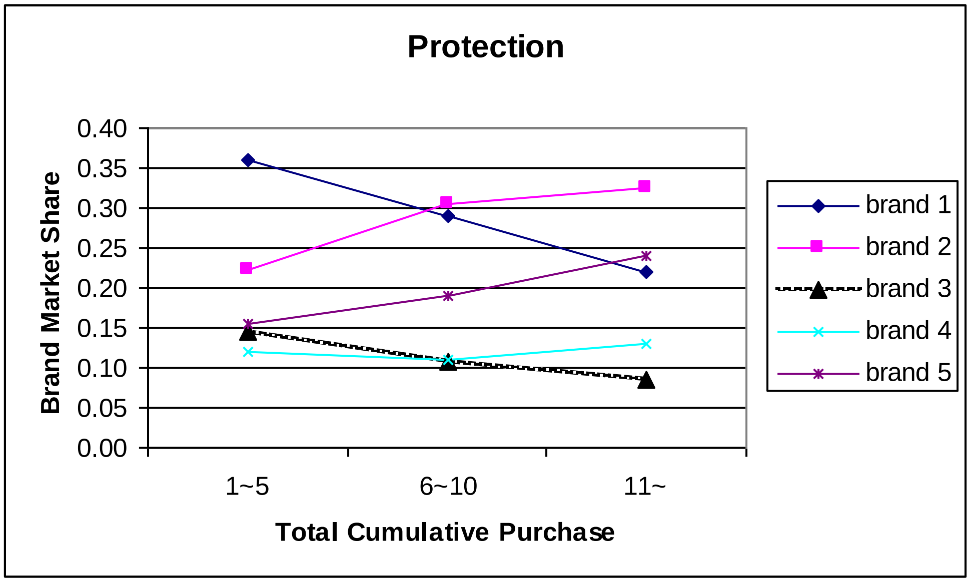
	Carabiner	Protection	Harness
Brand 1	35.17	27.61	16.42
Brand 2	11.78	30.99	17.64
Brand 3	10.87	10.17	24.92
Brand 4	21.23	10.81	15.45
Brand 5	20.94	20.42	25.57
	100%	100%	100%
Total Number of Observations	3,099	2,488	1,236

In order to determine if our raw data suggests brand preference evolution over time, we plotted brand market shares against the number of cumulative purchases made by customers across each of the three subcategories (Figure 1-3). Figure 1 illustrates the case of carabiners. The graph indicates that over the course of the customers' purchase histories, there is either an upward or downward trend in the market share of each of the five major brands of carabiners; as the number of customers' cumulative purchases increases, the market shares of Brands 1 and 2 increase, and the market shares of Brands 3, 4, and 5 decrease.

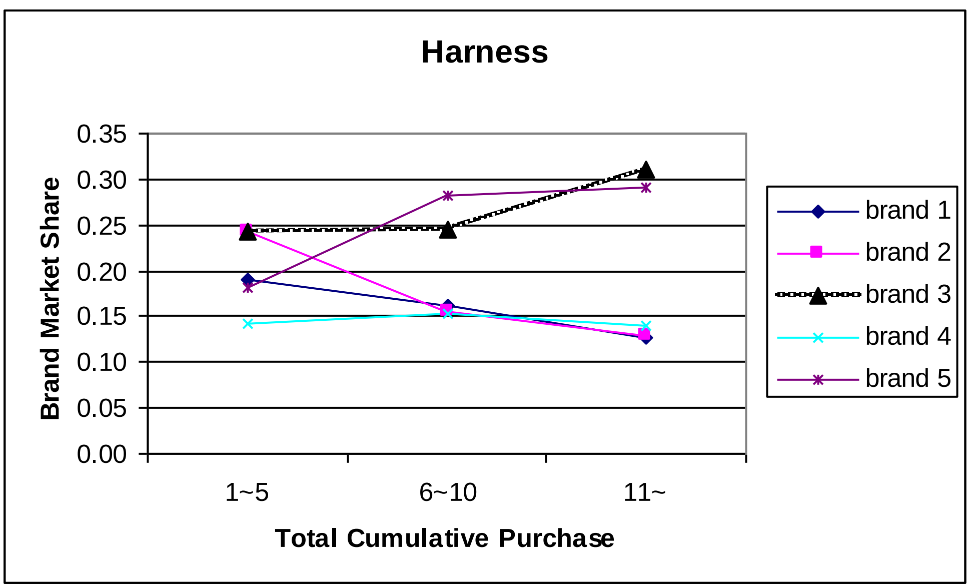
[Figure 1] Dynamics in Choice Behavior for Carabiner



[Figure 2] Dynamics in Choice Behavior for Protection



[Figure 3] Dynamics in Choice Behavior for Harnesses



For subcategory of protection, the market shares of Brands 1 and 3 decrease and the market shares of Brands 2 and 5 increase as the number of cumulative purchases increases. (See Figure 2.) The market share of Brand 4 does not change much over the customers' purchase history. For the harness subcategory, the market shares of Brand 1 and 2 decrease as the number of cumulative purchases increases. (See Figure 3.) Brands 3 and 5 are preferred as the number of cumulative purchases increases. The market share of Brand 4 stays stable.

3. Model

Utility is state dependent. In particular, consumers' sensitivity parameters are state specific. We hypothesize that consumer h's indirect utility for brand j in category c at time t is a function of intrinsic brand preference, price, and the last brand purchased (LBP). LBT is included to capture brand loyalty or the habitual behavior of customers (Heilmann et al. 2000). This variable represents state dependency, or inertia, and it changes over time. For example, consumer loyalty to one specific brand may grow or diminish through time, or consumers may seek increasing variety as they gain more experience in certain product categories. The state dependent choice probability is:

$$U_{hcjt} = \beta_{0,hcj}^{(s)} + \beta_{1,hc}^{(s)} P_{hcjt} + \beta_{2,hc}^{(s)} LBP_{hcjt} + \varepsilon_{hcjt}$$

$$LBP_{hcjt} = 1 \text{ if consumer h purchased brand j on t-1}$$
$$\qquad 0 \text{ otherwise}$$

where s=1,…, S denoting the consumer's given state. We have M product categories, c=1,.., M.

Since the state is consumer specific, we do not require the subscript h once the state is given. This is intuitive as HMM is itself a finite mixture model. If the state is fixed over time, the model is the usual latent class model that captures consumer heterogeneity by using segment specific coefficients (Kamakura and Russell 1989).[10] For identification purposes, we include one brand dummy fixed at zero. The state specific parameter vector is given by

$$\beta_c^{(s)} = (\beta_{0,c,1}^{(s)}, \ldots, \beta_{0,c,J_c-1}^{(s)}, \beta_{1,c}^{(s)}, \beta_{2,c}^{(s)})'$$

where J_c is the number of brands in category c. So we have (J_c+1)* S sensitivity parameters (i.e., J-1 intrinsic brand preferences, 1 price sensitivity, and 1 loyalty parameter for each state).

[10] HMM, like typical latent class models, classifies consumers into a finite set of states (or segments) based on their sensitivities to marketing variables. Unlike latent class models, however, HMM allows for dynamic switching of segment membership in the latent states through a Markov process. In other words, a latent class model can be seen as a HMM in which the transition matrix is an identity matrix.

In the HMM, states follow a Markov chain with a transition probability matrix π_{ht} that is consumer and time specific.[11]

The probability that consumer h is in state s at time t is given by

$$\phi_{hs,t} = \sum_{r=1}^{S} \phi_{hr,t-1} \pi_{hrst}$$

where π_{hrst} denotes the transition probability that the state of consumer h changes from r at t-1 to s at t. Specifically, we model this as:

$$\pi_{hrst} = \frac{\exp(\alpha_{h1r} + \sum_{c=1}^{M} \alpha_{h2r,c} E_{hc,t-1})}{1 + \exp(\alpha_{h1r} + \sum_{c=1}^{M} \alpha_{h2r,c} E_{hc,t-1})}, \text{ if } s = r+1 \text{ and } r < S$$

$$\pi_{hrst} = \frac{1}{1 + \exp(\alpha_{h1r} + \sum_{c=1}^{M} \alpha_{h2r,c} E_{hc,t-1})}, \text{ if } s = r \text{ and } r < S$$

$$\pi_{hrst} = 1, \text{ if } r = s = S$$

$$\pi_{hrst} = 0, \text{ otherwise}$$

where E denotes the consumer's cumulative experience level for each category. To capture cumulative experience, we use an objective measure based on consumer behavior. For each customer *h* in time period *t*, category experience is calculated as the number of cumulative purchases made by the customer in a particular subcategory of rock climbing equipment. Important to note is that our model does not allow for consumer to go back to a lower state[12]. Also, consumers may grow in terms of their knowledge and skills, but growth for individuals may occur at different rates.[13] Note that the chain is non-homogeneous in which the Markovian transitions are a function of time-

[11] We are using category specific α_h to set the individual specific parameters of the transition probabilities. A seemingly simpler option would be to use individual specific α_h with the cumulative experience $E_{h,t-1}$ calculated across all categories. This, however, would prevent us from addressing one of our research questions, what is the most representative category that can closely indicate consumer evolution? The category specific α_h's work as weight parameters which can inform us the relative importance of each category in switching up the consumers to the next experience state.

[12] Our model does not allow a consumer to go back to a lower state. The customer can stay at the state where she was, but can not move down. We assume the knowledge or skills acquired by a customer through learning can not be lost through forgetting. Learning is a process that depends on experience and leads to long-term changes in behavior potential (Ellis 1965). A future extension to this model could be using a more flexible transition structure. Inactivity for a long period of time may not result in a consumer's depreciation of knowledge, however, when consumers have not done rock climbing for a while, some knowledge may become irrelevant due to new advancements or technological improvement of products considering that sporting goods category is known for its vigorous R&D activities.

[13] And we do not have an exit state. If we needed to model purchase incidence, it would be appropriate to have an exit state.

varying covariates (Netzer, Lattin, and Srinivasan 2007, Hughes and Guttorp 1994). In addition, our model uses a non-stationary HMM (Netzer, Lattin, and Srinivasan 2007).

We need to model the initial probabilities (the probability that consumer h is in state s at time 0) as well. The initial probability is modeled by a logit probability

$$\phi_{hs,0} = \frac{\exp(\alpha_{h0s})}{\sum_{r=1}^{S} \exp(\alpha_{h0r})}$$

Again, we need to fix α_{h0S}=0 for identification purpose.

The enitre model is estimated using Markov Chain Monte Carlo (MCMC) hierarchical Bayes procedure. (For details see Appendix.)

4. Results

In this section, we present the results from the HMM model presented in section 3, and compare the fit and performance of our model with other potential benchmark models. We also discuss how consumer preferences vary over time and across individuals, and what drives these differences.

4.1 HMM

4.1.1 Model with Heterogeneity in Initial State Probability

In order to choose the number of states in our model, we use the marginal likelihood, or the posterior marginal density, as a model selection criterion following Newton and Raftery (1994). The marginal likelihood is calculated using a harmonic mean of the individual likelihoods from the posterior distribution.[14] [Table 4] reports the log marginal likelihoods for models with a different number of states. Based on the log marginal likelihood, the model with three states best fits the data.

[Table 4] The Number of States in HMM

Number of States	1	2	3	4
Log Marginal Likelihood	-10244	-10077	**-10006**	-10099

[14] For more applications of the marginal likelihood to HMM as a model evaluation criterion, see Montgomery et al. (2004), Lachaab et al. (2006), and Netzer, Lattin, and Srinivasan (2007).

[Table 5] reports the parameter estimates from the three-state model. From the results we notice that the intrinsic brand preference, price sensitivity, and last brand purchased parameters differ for each state. Some brands are preferred by beginners and others by more advanced consumers. In State 1, price parameters are positive. It seems that the role of price as a signal of quality is important in State 1. Since consumers in the beginner state have not yet acquired very much knowledge, they may simply choose the expensive brands. In the second state, price parameters are all negative. For the protection and the harness categories, State 2 consumers demonstrate the greatest price sensitivity. LBP in the protection category is significantly negative, perhaps indicative of variety seeking among consumers. LBP in the harness category is positive. State 3 consumers also exhibit positive LBP effects. Customers become more loyal or state-dependent in this state. Price sensitivity is highest among all states for the carabiner category.

[Table 5] 3-State Model Estimates

	State 1		State 2		State 3	
	average	std	average	std	average	std
Carabiner						
Brand1	**2.13**	0.16	**-0.73**	0.09	**0.08**	0.08
Brand2	**0.70**	0.16	**-1.44**	0.20	**-0.55**	0.09
Brand3	**1.92**	0.11	**-1.27**	0.12	**-1.85**	0.11
Brand4	**1.60**	0.10	**-0.21**	0.08	**-0.55**	0.10
LBP	**0.86**	0.13	-0.09	0.10	**0.37**	0.06
Price	**3.03**	0.09	**-4.96**	0.18	**-9.12**	0.13
Protection						
Brand1	**1.55**	0.14	**-2.73**	0.14	**-1.14**	0.10
Brand2	**0.56**	0.27	**1.41**	0.13	**0.55**	0.07
Brand3	**1.51**	0.12	**-3.81**	0.13	**-2.45**	0.18
Brand4	**1.19**	0.16	**-3.40**	0.07	**-1.55**	0.13
LBP	**1.84**	0.12	**-0.13**	0.09	**0.57**	0.08
Price	**1.93**	0.15	**-9.37**	0.14	**-2.11**	0.16
Harness						
Brand1	**2.74**	0.24	**-1.15**	0.10	**-0.48**	0.16
Brand2	**3.07**	0.10	**-0.48**	0.09	**-1.08**	0.08
Brand3	**2.18**	0.13	**-0.45**	0.08	0.03	0.12
Brand4	**3.77**	0.15	**-1.56**	0.15	**-0.52**	0.18
LBP	**0.73**	0.17	**0.64**	0.10	**0.88**	0.10
Price	**3.37**	0.28	**-2.42**	0.19	**0.87**	0.17

[Table 6] shows initial state probability. On average, 18% of customers belong to the beginner state, 68% to the intermediate state, and 14% to the more advanced state. The results indicate that the initial probability of State 2 is the greatest, and suggests that customers are likely to start their patronage of the retailer as intermediate rock climbers rather than true beginners. It is also possible that at the beginning of our data period, customers had already accumulated a degree of experience or knowledge. Since our data is left-truncated, we observe 68% of consumers in State 2 at the start of the data period.

[Table 6] Initial State Probability

	State 1		State 2		State 3	
	average	std	average	std	average	std
Initial Probability Parameter	**0.21**	0.16	**1.55**	0.05	0.00	Fixed
in terms of probabilities	0.18		0.68		0.14	

The parameter estimates for the transition matrix reported in [Table 8] show that the harness category is more indicative than other categories of a change from State 1 to State 2. In other words, the more harnesses a customer purchases , the more likely that customer will transition to a more advanced state. The intercept terms (α_{h1}) act like state specific fixed effects. These terms control for state specific unobserved effects that may exist. Also, the intercept terms represent stickiness to the states.

[Table 7] Transition Probability

	State1 to State2		State2 to State3	
	average	std	average	std
Intercept	**0.85**	0.07	**-5.75**	0.08
Carabiner	**-0.62**	0.08	**2.47**	0.18
Protection	**0.60**	0.07	**2.57**	0.21
Harness	**3.51**	0.11	**2.57**	0.14

4.1.2 Model with Fixed Initial State Probability

Next, we estimate the model with the initial state fixed at State 1. In this model, all consumers are assumed to start from the beginner state.[15] The heterogeneous parameters are those for transition probabilities only. Table 8 shows the estimation results from the 3-state model with fixed initial probability at State 1.

[Table 8] 3-State Model Estimates with Fixed Initial State

	State 1		State 2		State 3	
	average	std	average	average	std	average
Carabiner						
Brand1	**0.23**	0.07	**-0.54**	0.13	**0.49**	0.05
Brand2	**-1.15**	0.10	**-1.50**	0.10	**-0.49**	0.06
Brand3	**-0.41**	0.12	**-0.97**	0.19	**-0.68**	0.07
Brand4	**0.35**	0.11	0.04	0.14	-0.05	0.06
LBP	-0.09	0.13	**-0.36**	0.20	**0.40**	0.04
Price	**1.71**	0.11	**0.93**	0.12	**-0.92**	0.14
Protection						
Brand1	**1.50**	0.08	**0.94**	0.09	**-0.37**	0.07
Brand2	**0.44**	0.26	**0.12**	0.12	**0.47**	0.06
Brand3	**0.76**	0.08	-0.02	0.14	**-1.30**	0.09
Brand4	**0.43**	0.06	**-0.33**	0.18	**-1.02**	0.05
LBP	**-1.00**	0.06	**1.04**	0.07	**0.62**	0.04
Price	**0.72**	0.09	**-0.71**	0.16	**-1.34**	0.17
Harness						
Brand1	**0.30**	0.12	**0.33**	0.07	**-0.46**	0.18
Brand2	**0.55**	0.08	**0.28**	0.06	**-0.44**	0.06
Brand3	**0.41**	0.09	**0.39**	0.07	-0.05	0.06
Brand4	**-0.40**	0.18	**-0.25**	0.07	**-0.38**	0.07
LBP	-0.08	0.14	**0.44**	0.14	**0.83**	0.06
Price	**1.49**	0.19	**0.52**	0.05	-0.01	0.14

The coefficients in the utility functions prove to be consistent across categories. First, across all categories, State 3 customers exhibit "brand loyalty" while State 1 customers exhibit "variety seeking" behaviors. This indicates that customers are more likely to try different brands in the low states. In the higher states, consumers tend to buy brands that they previously purchased. Second, for all categories, State 1 is associated

[15] Even though we assume all consumers start as beginners, their rate of progress is heterogeneous. All consumers begin with easy tasks, and if their actual skills and knowledge levels are more advanced, then they will quickly transition to the next state.

with positive price sensitivity, and State 3 with negative sensitivity. This result is the opposite of the findings of Heilman et al. (2000). Our results indicate that if consumers are familiar with the product category, they are more price sensitive than when they are not familiar with the category. This observation should be particularly apparent for an activity like rock climbing where the stakes are so high. Certainly, we would not expect low-state climbers who do not have great knowledge of gear to compromise their safety by choosing cheaper brands of unknown quality or reliability.

In this model specification, all the (mean) coefficients in the transition probability are positive (See [Table 9]). More experience corresponds to greater probabilities among higher states.

[Table 9] Transition Probability with Fixed Initial State

	State1 to State2		State2 to State3	
	std	average	std	average
Intercept	**3.79**	0.20	**1.70**	0.14
Carabiner	**2.74**	0.16	**0.70**	0.20
Protection	**1.61**	0.16	**5.52**	0.38
Harness	**4.72**	0.28	**1.23**	0.09

4.2 Benchmark Model

We compare our HMM model with Heilman, Bowman, and Wright’s approach (HBW), and demonstrate the advantages and additional insights provided by our proposal.

HBW model

In order to describe the evolution of brand preference over time, Heilman et al. (2000) develop a model that incorporates the interaction between brand intercept, price, and LBP and customers’ cumulative purchases. In the HBW model, the consumer experience enters as covariates in the utility function. Though HBW accounts for time dynamics only in an ad hoc way, HMM takes a more behaviorally structured approach through the inclusion of a transition matrix, the data revealing the number of states and transition probabilities between the states.

The details of the HBW model are as follows;

Probability of consumer *h* choosing brand *j* on purchasing occasion *t* is modeled as logit:

$$P_{hjt} = \exp(U_{hjt})/(1+\sum_{j=1}^{J}\exp(U_{hjt}))$$

The utility function of consumer h is time dependent;

$$U_{hjt} = \alpha_{hj}(t) + \beta_h(t)x_{hjt} + \varepsilon_{hjt}$$

where a_j is the intrinsic utility of brand j for consumer h and x_{jt} is a vector of marketing variables, which includes all the covariates used in the HMM in our case.

For each customer h in time period t, cumulative experience is measured as the number of cumulative purchases made by the customer in a product category, which is captured in the variable CumPurch. The intrinsic brand utility and the vector of parameters for the marketing variables then become;

$$\alpha_{hj}(t) = \alpha_{0hj} + \alpha_{1hj}\ln(\mathrm{CumPurch}_{ht})$$
$$\beta_h(t)\, x_{hjt} = \beta_{0h}x_{hjt} + \beta_{1h}\ln(\mathrm{CumPurch}_{ht})x_{hjt}$$

In order to capture the brand loyalty or state dependency of customers, HBW includes a variable for the last brand purchased (LBP). The impact of LBP can change over time as the consumers gain more experience in product categories, i.e., consumers can become more loyal to one specific brand or they can become variety seeking. To capture these dynamics, HBW includes the interaction terms between LBP and cumulative purchases.

$$\beta_{2h}(t)\mathrm{LBP}_{hjt} = \beta_{02h}\mathrm{LBP}_{hjt} + \beta_{12h}\ln(\mathrm{CumPurch}_{ht})\mathrm{LBP}_{hjt}$$

The final model is estimated using a latent class heterogeneity structure;

$$U_{hjst} = \alpha_{0hjs} + \alpha_{1hjs}\ln(\mathrm{CumPurch}_{ht}) + \beta_{01hs}\mathrm{Price}_{hjt} + \beta_{11hs}\ln(\mathrm{CumPurch}_{ht})\mathrm{Price}_{hjt}$$
$$+ \beta_{02hs}\mathrm{LBP}_{hjt} + \beta_{12hs}\ln(\mathrm{CumPurch}_{ht})\mathrm{LBP}_{hjt} + \varepsilon_{hjt}$$

where s denotes the customer's segment.

The HBW model was estimated using the maximum likelihood procedure. [Table 10] reports Bayesian Information Criterion (BIC) as model comparison criteria for a simple multi-category latent class model in which segments are characterized by brand preference, price responsiveness, and the effect of LBP in multiple categories. The BIC indicates that the one segment model (homogeneous market) is preferred over two- or three- segment models. The parameter estimates of the one segment latent class model are shown in [Table 11]. [Table 12] shows that the one segment model best fits when using Heilman et al.'s model specification as well. The parameter estimates for the HBW model are reported in [Table 13].

[Table 10] Simple Multi-category Latent Class Model

Simple Latent Class Models Included variable - Brand Dummies, Price, and LBP			
Number of segments	1	2	3
Log Likelihood	10203.02	10122.16	10070.88
Number of observations	6823	6823	6823
Number of parameters	18	37	56
BIC	20564.94	20570.96	20636.13

[Table 11] The parameter estimates in simple Multi-category Latent Class Model

		Segment 1	
		Estimate	SE
Carabiner	BrandDummy1	-0.413	0.121
	BrandDummy2	-1.026	0.077
	BrandDummy3	-1.539	0.130
	BrandDummy4	-0.438	0.075
	LBP	0.412	0.038
	Price	-0.107	0.013
Protection	BrandDummy1	-0.486	0.033
	BrandDummy2	0.565	0.022
	BrandDummy3	-1.416	0.043
	BrandDummy4	-1.132	0.034
	LBP	0.616	0.036
	Price	-0.018	0.001
Harness	BrandDummy1	-0.193	0.024
	BrandDummy2	-0.229	0.062
	BrandDummy3	0.106	0.035
	BrandDummy4	-0.240	0.022
	LBP	0.884	0.034

Price	0.005	0.002

[Table 12] Model Selection for Heilman et al's Latent Class Model

Heilman et al's Model Included variable - Brand Dummies, Price, and LBP with interaction terms with cumulative purchases			
Number of segments	1	2	3
Log likelihood	10064.03	9982.61	9935.07
Number of observations	6823	6823	6823
Number of parameters	36	73	110
BIC	20445.86	20609.67	20841.22

[Table 13] The parameter estimates in homogeneous Heilman et al.'s model

		Est.	S.E.
Carabiner	BrandDummy1	-0.84	0.09
	BrandDummy2	-2.05	0.13
	BrandDummy3	-1.20	0.08
	BrandDummy4	0.03	0.11
	ln(Cumpurch) X BrandDummy1	0.65	0.05
	ln(Cumpurch) X BrandDummy2	0.83	0.06
	ln(Cumpurch) X BrandDummy3	0.24	0.04
	ln(Cumpurch) X BrandDummy4	-0.08	0.06
	LBP	-0.66	0.05
	ln(Cumpurch) X LBP	0.57	0.03
	Price	-0.40	0.04
	ln(Cumpurch) X Price	-1.20	0.11
Protection	BrandDummy1	0.87	0.06
	BrandDummy2	0.75	0.09
	BrandDummy3	-0.10	0.03
	BrandDummy4	-0.23	0.04
	ln(Cumpurch) X BrandDummy1	-0.97	0.06
	ln(Cumpurch) X BrandDummy2	-0.09	0.05
	ln(Cumpurch) X BrandDummy3	-0.96	0.07
	ln(Cumpurch) X BrandDummy4	-0.65	0.06
	LBP	0.03	0.02
	ln(Cumpurch) X LBP	0.32	0.02
	Price	-1.40	0.12
	ln(Cumpurch) X Price	-0.49	0.06
Harness	BrandDummy1	0.15	0.03
	BrandDummy2	0.32	0.05
	BrandDummy3	-0.06	0.03
	BrandDummy4	-0.39	0.03
	ln(Cumpurch) X BrandDummy1	-0.48	0.04
	ln(Cumpurch) X BrandDummy2	-0.61	0.05
	ln(Cumpurch) X BrandDummy3	0.04	0.03
	ln(Cumpurch) X BrandDummy4	-0.05	0.03
	LBP	0.17	0.03
	ln(Cumpurch) X LBP	0.55	0.03
	Price	-0.23	0.04

ln(Cumpurch) X Price	0.23	0.03

We find our HMM model performs better than the benchmark HBW model in identifying the dynamics of consumers' preference evolution. The homogeneous (one segment) HBW model performs better than the simple multi-category latent class model (BIC for the homogeneous HBW model is smaller than that of simple latent class model with any number of segments), so inclusion of cumulative purchase interaction terms to capture time dynamics works well to explain the data. The most of the interaction parameters are significant. The benchmark HBW model, however, fails to distinguish zero-order cross-sectional heterogeneity from time dynamics. Results from the HBW model show that after controlling for the time varying effects using the cumulative purchase interaction terms, and state dependence using the last brand purchased, there is no significant cross-individual heterogeneity left. Unlike the HBW model, our HMM incorporates both heterogeneity and time dynamics in the transition matrix using a full random effect specification – i.e., the transition matrix parameters, α_h, are individual specific.

Table 14 summarizes our comparison of the performances of the HMM and the other benchmark models. We report -2 log marginal likelihood for the HMM (Bayesian estimation), and -2 log likelihood and BIC for the HBW and simple logit model (Maximum Likelihood Estimation).[16] The HMM fits the data better than the HBW and the simple logit model.

[Table 14] The Fit Comparison between the HMM and the HBW Model

	HMM with 3 states	Homogeneous HBW model	Homogeneous simple logit model
-2 log marginal likelihood	20012		
-2 log likelihood		20128	20406
BIC		20445	20564

5. Conclusion

Understanding how consumer preferences vary over time is an important issue in brand choice studies. In this paper, we find evidence of brand preference evolution with

[16] The Bayesian Information Criterion (BIC) is minus twice the Schwarz Criterion, which is an asymptotic approximation to the log marginal likelihood or the Bayesian posterior marginal distribution (Schwarz 1978).

consumer experience or category learning associated with rock climbing. Our HMM model provides a parsimonious way to capture consumer evolution when we have a relatively short purchase history, as typical for data from the sporting goods industry. The HMM accounts for individual heterogeneity using a hierarchical Bayes procedure, which allows us to incorporate temporal variation and cross-sectional heterogeneity simultaneously. We compare our results to benchmark models and find the HMM outperforms the static latent class model which captures state dependence as a covariate that is included in the utility function. The HMM is superior in separating out time varying effects from individual cross-sectional heterogeneity.

Our results show that when it comes to buying recreational or sports equipment, customers' purchase decisions are influenced by the extent of their experience or knowledge of the activity or sport. Some products are better suited for beginners and others for more advanced customers. Customers' initial fitness and knowledge levels vary when they take up a sport for the first time, and over time they learn about the products available and about the sport itself at different rates. When customers attempt rock climbing for the first time, some start with easy tasks, but others start with a higher level of difficulty due to their better physical conditioning, experience in related activities, or perhaps having significant prior knowledge of the sport. Certain customers graduate to more advanced climbing tasks quickly, and others stay with more elementary tasks over a relatively long time. Thus, we find that consumers are heterogeneous in their skill levels as they begin sport climbing, and this extends to their rate of progress in the sport.

A particular strength of our model is that preference parameters vary systematically over time, and this may lead to a clearer and strategically meaningful understanding of choice and preference evolution. The HMM we develop suggests an effective way for a firm to dynamically segment and manage life-long relationships with its customer base by offering individualized product portfolios based on the varying experience levels of the customers. Our research has also yielded interesting patterns of variation over time in brand intercepts, price parameters, and loyalty measures. We find that consumers appear increasingly price sensitive over the time span of our data. Thus, our findings also suggest that retailers can optimize their marketing activities by taking into account the varying price sensitivities of customers through time.

Our study opens up many areas for future research. Our model does not allow for consumers who may through time lapse back to a lower state, for example, due to inactivity, injury, or aging. A future extension to our study is to consider more flexible transition structures. It would also be interesting to model other sporting goods categories and compare the effect sizes across different sports/recreation activities. The latent state concept is an effective method to capture unobserved consumer experience and learning, but the latency of the states does not allow us to determine the complete set of factors that constitute the experience state. We assume that the state specific fixed effect (constant term) and the weighted average of the cumulative purchases made in three gear categories are effective proxies for consumer experience. Future research may address this limitation through controlled experimentation.

References

Ainslie, A. and P.E. Rossi, 1998, "Similarities in Choice Behavior Across Multiple Categories," *Marketing Science*, 17, 2, 91-106

Anderson, Eric T., and Duncan I. Simester, 2004 "Long-Run Effects of Promotion Depth on New Versus Established Customers: Three Field Studies," *Marketing Science*, Vol. 23, No. 1, Winter, pp. 4-20.

Brangule-Vlagsma, Kristine, Rik Pieters, and Michel Wedel, 2002, "The dynamics of value segments: modeling framework and empirical illustration," *International Journal of Research in Marketing*, Vol. 19, No, 3, 267-285

Campbell, Margaret C. and Ronald C. Goodstein, 2001, "The Moderating Effect of Perceived Risk on Consumers' Evaluations of Product Incongruity: Preference for the Norm," *Journal of Consumer Research*, 28 (December), 439-449

Chib, Siddhartha, 1996, "Calculating Posterior Distributions and Model Estimates in Markov Mixture Models," *Journal of Econometrics*, 75, 79-97

Du, Rex Y. and Wagner A. Kamakura (2006), "Household Life Cycles and Lifestyles in the United States," *Journal of Marketing Research*, 43 (1), 121-32

Ellis, H., 1965, *The Transfer of Learning*, Macmillan Company, New York.

Erdem, 1998, "An Empirical Analysis of Umbrella Branding," *Journal of Marketing Research*, Vol. 35, August, 339-351.

Fitch, Stephane, 2005, "Uphill Battle," *Forbes*. Apr 25, 2005. Vol.175, No. 9, p. 62

Heilman, Carrie M., Douglas Bowman, and Gordon P. Wright, 2000, "The Evolution of Brand Preferences and Choice Behaviors of Consumers New to a Market," *Journal of Marketing Research*, Vol. 37, May, 139-155

Hoeffler S, D. Ariely, and P. West, 2006, "Path dependent preferences: The role of early experience and biased search in preference development," *Organizational Behavior and Human Decision Processes*, 101 (2): 215-229

Hoeffler, S. and D. Ariely, 1999, "Constructing Stable Preferences: A Look into Dimensions of Experience and Their Impact on Preference Stability," *Journal of Consumer Psychology*, Vol. 8, No. 2, 113-139.

Hughes, James P. and Peter Guttorp, 1994, "A Class of Stochastic Models for Relating Synoptic Atmospheric Patterns to Regional Hydrologic Phenomena," *Water Resources Research*, 30 (5), 1535-1546.

Kamakura, Wagner A. and Gary J. Russell, 1989, "A Probabilistic Choice Model for Market Segmentation and Elasticity Structure," *Journal of Marketing Research*, 26 (4), 379-390.

Lachaab, Mohamed, Asim Ansari, Kamel Jedidi, and Abdelwahed Trabelsi, 2006, "Modeling Preference Evolution in Discrete Choice Models: A Bayesian State-Space Approach," *Quantitative Marketing and Economics*, 4, 57-81

Liechty, John, Rik Pieters, and Michel Wedel, 2003, "Global and Local Covert Visual Attention: Evidence from a Bayesian Hidden Markov Model," *Psychometrika*, 68 (December), 519-41.

MacDonald, Iain L. and Walter Zucchini, 1997, *Hidden Markov and Other Models for Discrete-valued Time Series*, London: Chapman and Hall.

Montgomery, Alan L., Shibo Li, Kannan Srinivasan, John C. Liechty, 2004, "Modeling online browsing and path analysis using clickstream data," *Marketing Science*, 23 (4), 579-86

Moon, Sangkil, Wagner A. Kamakura, and Johannes Ledolter, 2007, "Estimating Promotion Response When Competitive Promotions are Unobservable," *Journal of Marketing Research*, Vol. 44, No. 3, 503-515

Netzer, Oded, James M. Lattin, and V. Srinivasan, 2007, "A Hidden Markov Model of Customer Relationship Dynamics," Working paper.

Newton, Michael A. and Adrian E. Raftery, 1994, "Approximate Bayesian Inference by the Weighted Likelihood Bootstrap," *Journal of the Royal Statistical Society*, B, 56 (1), 3-48.

Rabiner, L.R., 1989, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," *Proceedings of IEEE*, **77** (2), 257–285.

Schwarz, Gideon, 1978, "Estimating the Dimension of a Model," *The Annals of Statistics*, 6 (2), 461-464.

Scott, Steven L., 2002, “Bayesian Methods for Hidden Markov Models: Recursive Computing in the 21st Century,” *Journal of the American Statistical Association*, Vol. 97, No. 457, 337-351

West, M., and J. Harrison, 1997, *Bayesian Forecasting and Dynamic Models*, New York, Springer.

Appendix

1. Estimation Details

1.1 Parameters to be estimated

In terms of the number of parameters, we have the following set of parameters for each consumer:

α_{h0r} , r=1,… S-1

α_{h1r} , r=1,… S-1

$\alpha_{h2r,c}$, r=1,… S-1, c=1,…,M

So we have (M+2)*(S-1) individual specific parameters. For example, when M=3 and S=3, we have 10 individual parameters related to consumer evolution.

Note that β is common across consumers and that α_h is consumer specific. The dimension of β is $S * \sum_{c=1}^{M}(J_c + 1)$ and that of α_h is (M+2)*(S-1). Since α_h is consumer specific, we model it as a continuous random variable normally distributed across consumers. Specifically,

$$\alpha_h \sim N(a, V_\alpha) .$$

So, what we need to estimate is (β, a, V_α).

1.2 Priors

$$\beta \sim N(b_0, V_{\beta 0})$$

$$a \sim N(a_0, V_0)$$

$$V_\alpha \sim IW(f_0, G_0^{-1})$$

1.3 MCMC Algorithm

The direct Metropolis-Hastings procedure would average over states when it comes to the likelihood calculation at the given value of parameters. As well-documented in Scott (2002), there are several reasons to prefer sampling states instead of integrating them out. We utilize the approach developed by Chib (1996) in which the unobserved states are simulated. It can also be considered a data-augmentation approach.

So the MCMC procedure is a hybrid one in which Gibbs and M-H are used together as follows.

(1) Data augmentation: drawing states $S \mid \beta, \{\alpha_h\}, \mathrm{Data}, b_0, V_{\beta 0}$

First, we need to specify the conditional distribution of states given data and parameter values. We use the following notations: $S_t = \{s_1, \ldots, s_t\}$ and $S^{t+1} = \{s_{t+1}, \ldots, s_T\}$ where T is the length of the time series. Let D_T be the data. The joint distribution of states can be represented by the products of conditionals as follows:

$$p(S_T \mid \mathrm{Data}, \beta, \alpha) = p(s_T \mid D_T, \beta, \alpha)\, p(s_{T-1} \mid S^T, D_T, \beta, \alpha) \ldots p(s_1 \mid S^2, D_T, \beta, \alpha).$$

Using Bayes theorem

$$\begin{aligned} p(s_t \mid S^{t+1}, D_T, \beta, \alpha) &= p(s_t \mid S^{t+1}, D_t, D^{t+1}, \beta, \alpha) \\ &\propto p(s_t \mid D_t, \beta, \alpha) * p(S^{t+1}, D^{t+1} \mid s_t, D_t, \beta, \alpha) \\ &\propto p(s_t \mid D_t, \beta, \alpha) * p(s_{t+1} \mid s_t, D_t, \beta, \alpha)\, p(S^{t+2}, D^{t+1} \mid s_t, D_t, \beta, \alpha) \\ &\propto p(s_t \mid D_t, \beta, \alpha) * p(s_{t+1} \mid s_t, \alpha) \end{aligned}.$$

The last line of the above expression is due to (i) data is not necessary in determining state once the last state is given and (ii) Markov property where only the last time period matters in determining the current states.

In the above expression the second term (the transition probability) is given by the model. What is $p(s_t \mid D_t, \beta, \alpha)$? Note that

$$\begin{aligned} p(s_t \mid D_{t-1}, d_t, \beta, \alpha) &\propto p(s_t \mid D_{t-1}, \beta, \alpha) * p(d_t \mid D_{t-1}, s_t, \beta, \alpha) \\ &\propto p(s_t \mid D_{t-1}, \beta, \alpha) * p(d_t \mid s_t, \beta, \alpha) \end{aligned}$$

The second line is due to the fact that the observations are linked only through states in HMMs. Chib (1996) explains how to compute the probability using forward recursions by alternating the prediction step and the update step.

Prediction Step:

$$p(s_t \mid D_{t-1}, \beta, \alpha) = \sum_{k=1}^{S} p(s_t \mid s_{t-1} = k, \alpha) * p(s_{t-1} = k \mid D_{t-1}, \beta, \alpha)$$

where S is the number of states in the model. In order to determine $p(s_t \mid D_{t-1}, \beta, \alpha)$, we need to know $p(s_{t-1} \mid D_{t-1}, \beta, \alpha)$. Chib (1996) sets at t=1,

$p(s_1 | D_0, \beta, \alpha)$ to be the stationary distribution of the Markov transition matrix. So the prediction step is not required for t=1 in his approach. However, our model is nonstationary and nonhomogenous Markov model, with an initial probability distribution corresponding to $p(s_1 | D_0, \beta, \alpha)$.

Update Step:

$$\begin{aligned} p(s_t | D_t, \beta, \alpha) &\propto p(s_t | D_{t-1}, \beta, \alpha) * p(d_t | D_{t-1}, s_t, \beta, \alpha) \\ &\propto p(s_t | D_{t-1}, \beta, \alpha) * p(d_t | D_{t-1}, \beta_{s_t}) \end{aligned}$$

After finishing the forward recursion, we can compute the joint density as the product of conditionals as follows:

$$p(S_T | D_T, \beta, \alpha) = p(s_T | D_T, \beta, \alpha) \ldots * p(s_t | D_T, S^{t+1}, \beta, \alpha) \ldots * p(s_1 | D_T, S^2, \beta, \alpha).$$

As the conditionals are already available, we can simply draw the state one by one using the conditionals. As in Chib (1996), we first prepare a T-by-S matrix F that contains $p(s_t | D_t, \beta, \alpha)$. Given the (t-1)th row F_{t-1}, the next row F_t is proportional to $(F'_{t-1}\pi_{t-1}) \bullet \lambda_t$ where π is the transition matrix , λ is a row vector consisting of $p(d_t | D_{t-1}, \beta_{s_t})$, and $\bullet$ is element-by-element multiplication operator. We use the matrix F to simulate the states. First, we draw s_T, the state of the last period, from $p(s_T | D_T, \beta, \alpha)$. Next, we can simulate $s_t | S^{t+1}, t = 1, \ldots, T-1$ from the following distribution:

$$p(s_t | D_T, S^{t+1}, \beta, \alpha) \propto p(s_t | D_t, \beta, \alpha) * p(s_{t+1} | s_t, \alpha).$$

(2) Draw $\beta | \{\alpha_h\}, \text{States}, \text{Data}, b_0, V_{\beta 0}$

$$\begin{aligned} & p(\beta | \{\alpha_h\}, \text{State}, \text{Data}, b_0, V_{\beta 0}) \\ & \propto N(b_0, V_{\beta 0}) * L(\text{Data} | \{\alpha_h\}, \beta_{\text{state}}) \\ & \propto \exp\left(-\frac{1}{2}(\beta - b_0)' V_{\beta 0}^{-1} (\beta - b_0)\right) * L(\text{Data} | \{\alpha_h\}, \beta_{\text{state}}) \end{aligned}$$

Note that the likelihood is simple once the states are given. We can simply use the state-specific parameter vector. The likelihood for an observation is given by

$$\prod_{j=1}^{J}\left(\frac{\exp(x_{jt}'\beta_{s_t})}{\sum_{k=1}^{J}\exp(x_{kt}'\beta_{s_t})}\right)^{d_{jt}}.$$

Since we do not have a closed form for this, we use Metropolis Hastings algorithm with a random walk chain. First, we draw $\beta_{new} = \beta_{current} + c * N(0, I)$. Then the acceptance probability is given by

$$q = \min\left\{\frac{\exp\left(-\frac{1}{2}(\beta_{new}-b_0)'V_{\beta 0}^{-1}(\beta_{new}-b_0)\right) * L(Data \mid \{\alpha_h\}, \beta_{new})}{\exp\left(-\frac{1}{2}(\beta_{current}-b_0)'V_{\beta 0}^{-1}(\beta_{current}-b_0)\right) * L(Data \mid \{\alpha_h\}, \beta_{current})}, 1\right\}$$

and

$$\beta_{next} = \begin{cases} \beta_{new}, \text{ with prob } q \\ \beta_{current}, \text{ with prob } 1-q \end{cases}$$

(3) Draw $\alpha_h \mid \beta, States, Data, a, V_\alpha$

$$\begin{aligned} p&\left(\alpha_h \mid Data, State, \beta, a, V_\alpha\right) \\ &\propto N(\alpha_h \mid a, V_\alpha) * L(State_h \mid \alpha_h) \\ &\propto \exp\left(-\frac{1}{2}(\alpha_h - a)'V_\alpha^{-1}(\alpha_h - a)\right) * L(State_h \mid \alpha_h) \end{aligned}$$

Note that the choice data are irrelevant for the transition probability parameters once the states are given. The state data tell us everything about α. Define a new variable $g_{hmlt} = 1$ if $s_{ht} = l$ and $s_{h,t+1} = m$, and 0 otherwise. So g_{hmlt} indicates whether a state transition from l to m occurs at time t. The joint density can be represented by the new variables as follows:

$$p(s_{h1}, ..., s_{hT}) = p(s_{h1}, g_{h..1}, g_{h..2}, ..., g_{h..T}).$$

The log likelihood is given by

$$\log p\left(s_{h1}\right) + \sum_{t=1}^{T-1}\sum_{l=1}^{S}\sum_{m=1}^{S} g_{hlmt} * \log(\pi_{hlmt})$$

where the likelihood of the first observation is the initial probability of the first state.

Again, no closed form for the posterior distribution is available. So we use M-H again. First, we draw $\alpha_{h,new} = \alpha_{h,current} + c * N(0, I)$. Then use the following acceptance probability to determine whether to accept the candidate:

$$q = \min\left\{ \frac{\exp\left(-\frac{1}{2}(\alpha_{h,new} - a)' V_\alpha^{-1} (\alpha_{h,new} - a) \right) * L(State_h \mid \alpha_{h,new})}{\exp\left(-\frac{1}{2}(\alpha_{h,current} - a)' V_\alpha^{-1} (\alpha_{h,current} - a) \right) * L(State_h \mid \alpha_{h,current})}, 1 \right\}$$

(4) Draw $a \mid \{\alpha_h\}, a_0, V_a$

$$p \quad a \mid \{\alpha_h\}, V_\alpha, a_0, V_0 \quad = N(a_n, V_n)$$

where $V_n = \quad HV_a^{-1} + V_0^{-1} \quad ^{-1}$ where H is the number of consumers in the data set.

and

$$a_n = V_n * \left(HV_\alpha^{-1} \frac{\sum_{h=1}^{H} \alpha_h}{H} + V_0^{-1} a_0 \right) = V_n * \quad V_\alpha^{-1} \sum_{h=1}^{H} \alpha_h + V_0^{-1} a_0$$

$$= \quad HV_\alpha^{-1} + V_0^{-1} \quad ^{-1} \quad V_\alpha^{-1} \sum_{h=1}^{H} \alpha_h + V_0^{-1} a_0$$

(5) Draw $V_\alpha \mid \{\alpha_h\}, a$

$$p \quad V_\alpha \mid \{\alpha_h\}, a \quad = IW\left(f_0 + H, G_0^{-1} + \sum_{h=1}^{H} (\alpha_h - a)(\alpha_h - a)' \right)$$

2. Structure of Consumer Heterogeneity

As we account for consumer heterogeneity in initial state probability and transition probability coefficients, here we report the covariance and the correlation matrix to show the structure of consumer heterogeneity in Table 15 and Table 16.

[Table 15] Covariance Matrix

covariance matrix	init(1)	init(2)	intp(1,2)	Cara(1,2)	Prot(1,2)	Harn(1,2)	intpt(2,3)	Cara(2,3)	Prot(2,3)	Harn(2,3)
initial Prob (State 1)	0.19									
initial Prob (State 2)	0.03	0.26								
intercept (1,2)	-0.04	-0.05	0.22							
Carabiner (1,2)	-0.06	-0.03	0.01	0.29						
Protection (1,2)	-0.01	0.11	0.11	-0.05	0.22					
Harness (1,2)	0.10	0.09	-0.11	0.00	0.03	0.34				
intercept (2,3)	-0.11	-0.18	0.19	0.11	-0.10	-0.24	0.56			
Carabiner (2,3)	-0.01	0.02	0.01	-0.09	0.14	0.01	-0.07	0.74		
Protection (2,3)	-0.03	-0.12	0.20	-0.03	-0.10	-0.22	0.23	0.10	0.53	
Harness (2,3)	-0.03	-0.04	0.01	0.06	-0.01	-0.03	0.04	-0.06	0.00	0.14

[Table 16] Correlation Matrix

correlation	init(1)	init(2)	intp(1,2)	Cara(1,2)	Prot(1,2)	Harn(1,2)	intpt(2,3)	Cara(2,3)	Prot(2,3)	Harn(2,3)
initial Prob (State 1)	1.00									
initial Prob (State 2)	0.15	1.00								
intercept (1,2)	-0.19	-0.22	1.00							
Carabiner (1,2)	-0.25	-0.12	0.06	1.00						
Protection (1,2)	-0.07	0.44	0.48	-0.18	1.00					
Harness (1,2)	0.39	0.30	-0.40	0.00	0.12	1.00				
intercept (2,3)	-0.34	-0.47	0.54	0.28	-0.28	-0.55	1.00			
Carabiner (2,3)	-0.02	0.05	0.02	-0.20	0.36	0.02	-0.10	1.00		
Protection (2,3)	-0.11	-0.31	0.59	-0.08	-0.31	-0.51	0.43	0.16	1.00	
Harness (2,3)	-0.18	-0.19	0.07	0.30	-0.06	-0.14	0.14	-0.18	0.02	1.00