Inertia and Variety Seeking in a Model of Brand-Purchase Timing

Pradeep K. Chintagunta
Graduate School of Business, University of Chicago, Chicago, Illinois 60637
pradeep.chintagunta@gsb.uchicago.edu

Abstract

Previous research on state dependence indicates that a brand's purchase probabilities vary over time and depend on the levels of inertia and variety seeking and on the identity of the previously purchased brand. Brand-choice probabilities obtained from models such as the logit and the probit are, however, fixed over time, conditional on the previous brand purchased and on the levels of marketing variables. Consequently, state dependence has largely been studied as a time-invariant phenomenon in brand-choice models, with the levels of inertia and variety seeking assumed to be constant over time. To account for the time-varying nature of state dependence would require a model in which brand-switching probabilities depend upon interpurchase times. One modeling framework that can account for this dependence is based on the hazard function approach.

The proposed approach works as follows. All other factors being equal, an inertial household purchasing a brand on a particular occasion is most likely to repurchase that brand on the next occasion. If the household switches, it will be to a brand located perceptually close, in attribute space, to the previously purchased brand. In other words, an inertial household has the highest switching hazard for the same origin and destination brands, with a progressively lower hazard rate for brands perceptually located farther and farther away from the origin brand. The amount by which the hazard is lowered depends upon the perceptual distance and the inertia level of the household. On the other hand, if the household is variety seeking, the most likely brand purchased would be a brand located farthest away from the previously purchased brand in attribute space. In other words, the hazard rate of repurchase is the lowest, with the rate increasing with the distance of the destination brand from the origin brand and the level of that household's variety-seeking tendency. The effects of inertia and variety seeking are, therefore, incorporated at the attribute level into a brand-purchase timing model. In doing so, we attempt to provide greater insight into the nature of state dependence in models of purchase timing.

Our model and estimation procedure will enable us to distinguish between households that are inertial and those that are variety prone. In addition to accounting for state dependence, the model also accounts for the effects of unobserved heterogeneity among households in their brand preferences and in their sensitivities to marketing activities. A majority of studies in marketing using the hazard function approach to investigate purchase timing have not accounted for heterogeneity in marketing-mix effects. The study integrates recent methods that incorporate the effects of inertia and variety seeking in brand-choice models with a semi-Markov model of purchase timing and brand switching. The proposed model enables us to (1) infer market structure via a perceptual map for the sample households, and (2) investigate implications for the introduction of a line extension.

We provide empirical applications of the proposed method using three different household-level scanner panel data sets. We find that differing levels of inertia and variety seeking characterize the three data sets. The findings are consistent with prior beliefs regarding these categories. In addition, our results indicate that the nature of interbrand purchase timing behavior depends upon the extent of inertia or variety seeking in the data. We are also able to characterize the structure of the three product markets studied. This provides implications for interbrand rivalry in the market. Further, we demonstrate how the model and results can be used to predict the location of a line extension in the perceptual space of households. Finally, we obtain implications for the timing of brand promotions.
(*Hazard Models*; *Purchase Timing*; *Brand Choice*)

0732-2399/98/1703/0253$05.00

1. Introduction

State dependence in brand-choice models refers to the influence of previous purchases made by a household on that household's current purchase. Including a lagged purchase indicator (or a function of lagged purchases) in the indirect brand utility function typically operationalizes state dependence. A given household can have positive state dependence or *inertia* (previous purchase of brand *X* increases the probability of a current brand *X* purchase); or can have negative state dependence (previous purchase of brand *X* decreases the probability of a current brand *X* purchase), in other words, the household is *variety seeking*. From a managerial standpoint, it would be important to know whether a brand's consumers are inertial or variety prone. For example, finding that households are variety seekers may motivate managers to expand their product offerings, as they would like households to switch to their own brands. On the other hand, a preponderance of inertial households would shift emphasis to brand retention, as a switch away from the brand could lead to a sustained defection and the loss of the household as a customer.

The literature on variety seeking suggests that households seek variety along the attributes that characterize a brand (or object). Specifically, purchase of a brand results in a satiation of the attributes associated with that brand. Consequently, a household is less likely to purchase a brand with *those* attributes on the next purchase occasion (McAlister and Pesemier 1982). On the other hand, an inertial household will tend to purchase a brand that has similar attributes as those purchased on the previous occasion.

Marketing researchers (e.g., Trivedi, Bass, and Rao 1994) have found that a household's state dependence level varies over time. Specifically, a household's level of inertia may decline over time due to a diminishing ability to "recall" the last brand purchased (see for example, Roy, Chintagunta, and Haldar 1996). Hence, if the time between purchases is large, one might expect low levels of inertia for that household. Furthermore, as noted by McAlister (1982), a household's propensity for variety seeking would be highest immediately after brand purchase due to attribute satiation. Over time, the household becomes less satiated, thereby lowering its desire to seek variety. In this case, large interpurchase times could make brand switching less likely.

Taken together, previous research on state dependence indicates that a brand's purchase probabilities vary over time and depend on the levels of inertia and variety seeking and on the identity of the previously purchased brand. Brand-choice probabilities obtained from such models as the logit and the probit are, however, fixed over time conditional on the previous brand purchased and on the levels of marketing variables. Consequently, state dependence has largely been studied as a time-invariant phenomenon in brand-choice models, with the levels of inertia and variety seeking assumed to be constant over time (see Erdem 1996 for a recent illustration). To account for the time-varying nature of state dependence would require a model in which brand-switching probabilities depend upon interpurchase times (as described previously). One modeling framework that can account for both brand-choice decisions as well as interpurchase times of households is based on the hazard function approach (Jain and Vilcassim 1991).

Our objective in this paper is to incorporate the effects of attribute-based inertia and variety seeking into a brand-purchase timing model. In doing so, we attempt to provide greater insight into the nature of state dependence in models of purchase timing. Brand-switching probabilities in the proposed model would change over time for a household depending upon that household's level of inertia or variety seeking. Our model and estimation procedure will enable us to distinguish between households that are inertial and those that are variety prone. In addition to accounting for state dependence, the model also accounts for the effects of unobserved heterogeneity among households in their brand preferences and in their sensitivities to marketing activities. A majority of studies in marketing using the hazard function approach to investigate purchase timing have not accounted for heterogeneity in marketing-mix effects.

The proposed approach works as follows. All other factors being equal, an inertial household purchasing brand *i* on occasion *t* is most likely to repurchase brand *i* on the next occasion. If the household switches, it will be to a brand located perceptually close to brand *i*. In other words, an inertial household has the highest switching hazard for the same origin and destination brands, with a progressively lower hazard rate for brands perceptually located farther and farther away from the origin brand. The amount by which the hazard is lowered depends upon the *perceptual distance* and the *inertia level* of the household. On the other hand, if the household is variety seeking, the most likely brand purchased after the purchase of brand i would be a brand located farthest away from brand i in perceptual space. In other words, the hazard rate of repurchase is the lowest, with the rate increasing with the distance of the destination brand from the origin brand and the level of that household's variety-seeking tendency.

The perceptual distance between brands has been interpreted as a measure of interbrand competition (see for example, Elrod and Keane 1995). The ability of the model to account for such competition implies that one can uncover the market structure for a product market using the proposed approach. Previous researchers have attempted to uncover market structure by exploiting interbrand-purchase timing information (see for example, Fraser and Bradford 1983; Grover and Rao 1988). By adopting the proposed approach, however, one can obtain the perceptual locations of households after controlling for the effects of marketing variables, unobserved heterogeneity, and the effects of inertia and variety seeking. Hence, the study makes some progress with respect to the stream of literature on deriving market structure implications using purchase timing data.[1]

There are several interesting substantive issues that one can examine with the proposed methodology. For example, one can study the impact of repositioning an existing brand in the marketplace to a more "favorable location." Or, one can perform "what-if" scenario analyses with respect to the market structure and household behavior in terms of the levels of inertia and variety seeking. To illustrate this point, we provide a specific application of the methodology. Using the preference decomposition approach proposed by Fader and Hardie (1996), we demonstrate how one can predict the perceptual location of an imitative line extension using the model. Another issue of managerial interest pertains to the timing of brand promotions. If brand-switching probabilities vary over time, then managers can exploit this variation to time their promotions appropriately. Under a specific set of conditions we show how one can obtain implications for the timing of promotions.

The rest of this paper is organized as follows. In the next section we describe the model formulation. We then provide an empirical application of the proposed methodology to liquid and powdered laundry detergent data in two different geographic markets, as well as to a soft drinks submarket. The final section concludes.

2. Model Formulation

Vilcassim and Jain (1991) propose a semi-Markov hazard function model to examine the purchase timing and brand-switching behavior of households. The authors estimate the hazard rate of switching from brand A to brand B for each pair of brands in the category, including the case when $A = B$. The model implies, therefore, that there is a hazard associated with the exit from one state (brand A) and entry into another (brand B). We use this approach as a starting point for our model formulation.

As noted by Jain and Vilcassim (1991), the chosen specification for the hazard function must be flexible to allow for a variety of shapes of the function, most importantly, a monotonic increasing pattern or a decreasing pattern or a nonmonotonic pattern in which the hazard first increases and then declines. We investigated several functional forms with these properties, including the Box-Cox, inverse Gaussian, and the log-logistic, and concluded that on the three measures of model fit, predictive ability, and computational time, the log-logistic was the preferred specification. Hence,

[1]The market structure implications obtained by accounting for purchase timing and brand-choice decisions could differ from those obtained using information only on brand-choice decisions. The proposed model implies that brand-switching probabilities vary over time with the temporal dependence differing across brand pairs, unlike the case when using only brand-choice data where these probabilities are time invariant. As market structure is derived using brand-switching probabilities, the above difference can contribute to obtaining different implications from the two approaches.

for our purposes here, we choose the two-parameter log-logistic specification. The hazard for household i at time t of switching from brand j to brand k, $\lambda_{ijk}(.)$ is written as:

$$\lambda_{ijk}(t) = \frac{\gamma_{0i}\gamma_{1ijk}^{\gamma_{0i}}\ t^{\gamma_{0i}-1}}{1 + \gamma_{1ijk}^{\gamma_{0i}}\ t^{\gamma_{0i}}}, \tag{1}$$

where the two parameters are γ_{0i} and γ_{1ijk}, both of which are greater than zero. Hence, they can be expressed as: $\gamma_{0i} = \exp(\xi_{0i})$ and $\gamma_{1ijk} = \exp(\xi_{1ijk} + \mathbf{X}_{ijk}\theta_{2jk})$, where θ_{2jk} is a set of parameters to be estimated and $\mathbf{X}$ denotes the vector of covariates. Note that as these covariates directly impact the failure times, the above specification belongs to the class of accelerated failure time models.[2]

In the empirical analysis we find that making the θ_{2jk} parameters switch specific does not improve model fit (adjusting for the increase in parameters and for sample size). Hence, in the following, we set $\theta_{2jk} = \theta_2$. Note, however, that heterogeneity in these parameters across households is still accounted for in the estimation. Furthermore, the literature has suggested two possible operationalizations for $\mathbf{X}_{ijk}$. One is to set $\mathbf{X}_{ijk} = \mathbf{X}_{ik}$ and use the values of the covariates only for the brand purchased at time t. The other is to use a function of the covariates (such as the difference or the ratio) at times $t - 1$ and t that correspond to brands j and k, respectively. In our empirical analysis we tried both formulations and found that the former fit the data better. Hence, for the rest of the discussion, we assume $\mathbf{X}_{ijk} = \mathbf{X}_{ik}$.

Estimating a set of ξ_{1ijk} parameters for the above hazard function for each brand pair $j - k$ would result in a large number of parameters to be estimated. A more parsimonious approach would be to parameterize ξ_{1ijk} as follows:

$$\xi_{1ijk} = \delta_{ik} + \alpha_{ijk}. \tag{2}$$

δ_{ik} (in Equation (2)) is a brand (k)-specific parameter for household I that captures the relative rate of *repurchasing* brand k. If $\delta_{ij} > \delta_{ik}$, it implies that the repurchase hazard for brand j exceeds that for brand k. α_{ijk} is the parameter associated with a *switch* from brand j to brand k (for household i) with $\alpha_{ikk} = 0$ for all k. Hence, the parameter α_{ijk} measures the deviation of the baseline switching hazard from the baseline repurchase hazard and thereby accounts for *state dependence*.

Consider purchases at times $(t - 1)$ and t. Further, at time t brand k is purchased. State dependence exists if the brand purchased at time $(t - 1)$ influences that purchased at time t. It is clear from Equation (2) that $\lambda_{ijk}(.) \neq \lambda_{ilk}(.)$ for $j \neq l$ as long as $\alpha_{ijk} \neq \alpha_{ilk}$. Hence, current state is influenced by past state. It is also true that $\lambda_{ijk}(.) \neq \lambda_{ijl}(.)$ for $k \neq l$ as long as $\alpha_{ijk} \neq \alpha_{ijl}$. Thus, current state influences future states. In the context of hazard models, we have accounted for state dependence. Next, we discuss how can we separate the effects of inertia from those of variety seeking.

For an inertial household characterized by repurchasing, one would expect $\alpha_{ijk} < 0$ (for all $k \neq j$) as $\alpha_{ijj} = 0$. In other words, the hazard of repurchasing brand j exceeds the hazard of switching to brand k. Similarly, the variety-seeking tendency of the household would be reflected in $\alpha_{ijk} > 0$ (for $k \neq j$), i.e., a household is more likely to buy a brand different from brand j (the brand previously purchased). Households could differ on their extent of loyalty or variety seeking. Hence, one could use a random effects approach to estimate the distribution of α_{jk} across households i. This way, the nature of inertia and variety seeking across households can be studied.

As the number of brands increases, however, estimation of the model parameters would pose a problem. For example, with 10 brands, there would be 90 α_{jk} parameters. Allowing for a joint distribution (or even independent distributions) across these parameters is a daunting task. A more parsimonious approach, therefore, would be to impose a factor structure on the α_{jk} and rewriting (2) as follows:

$$\xi_{1ijk} = \delta_{ik} + \mu_i d_{jk},$$

where μ_i denotes household i's inertia (if $-1 < \mu_i < 0$) or variety-seeking (if $1 > \mu_i > 0$) tendency and $d_{jk} > 0$ is the perceptual distance between brands j and k. As the distance parameters are larger than zero, the sign

[2]Note that we could, alternatively, have made the other log-logistic parameter γ_{0i} a function of the marketing variables. However, interpretation of the effects would have been more complicated. In additional, we found that the model fit the data better under our current operationalization.

of the μ_i parameter reveals whether household i is inertial or variety seeking. Note that μ_i does not depend upon the identity of the brand. So given a μ_i for a household, a rank ordering of the d_{jk} reveals which two brands are considered most similar, etc.[3] While this is a simplification of the model in (1), it can be made even more parsimonious as follows: Let the perceptual space of households be restricted to be a (say) F-dimensional space (see, for example, Elrod and Keane 1995), then d_{jk} can be operationalized as the Euclidean distance between brands in this space (see Erdem 1996 for such a formulation in the context of brand-choice behavior).[4] Hence the above equation can be written as:

$$\xi_{1ijk} = \delta_{ik} + \sum_{f=1}^{F} \mu_{if} \, (l_{jf} - l_{kf})^2, \qquad (2')$$

where l_{jf} denotes the location of brand j along attribute (or dimension) f and the loyalty/variety-seeking parameter μ_{if} $(-1, 1)$ is made attribute specific as in McAlister and Pesemier (1982). The relative brand locations are then directly interpretable for their market structure implications. Note that we assume attributes to be continuous in nature. In several product categories, attributes are likely to be discrete. This is a limitation of the proposed formulation.[5]

As stated in the Introduction, a market structure methodology is useful in as much as it enables us to study interesting substantive problems. Here we illustrate an application of the framework.

Predicting Perceptual Locations of Line Extensions
In a recent paper, Fader and Hardie (1996) proposed a method to forecast the sales of imitative line extensions. They decompose the preference for an SKU (stock keeping unit) into the preferences for that SKU's attributes like brand name, size, form, formula, etc., when estimating the parameters of an SKU-level choice model using household data. We propose a similar method for predicting the perceptual location of a line extension. Consider a case where a "brand" is characterized by its brand name and size. The location parameters in our proposed model, l_{jm}, can be viewed as being driven by name and size. Note that this does not imply that the attributes of the derived perceptual map are name and size. What this means is that a brand's value along each attribute depends on the name and on the size of the brand.

Suppose that the data consist of purchases from a subset of all possible brand-size combinations such that all brands and all sizes are represented in the subset. Then once the model parameters are estimated, we can predict the locations of brand-size combinations that are not in the subset. This is possible because the parameters for the corresponding brands and sizes are now available. In a similar manner, the brand-specific parameters δ_j can also be decomposed into "brand" and "size" components.

Likelihood Function
Let the set of parameters to be estimated be denoted by $\Theta = (\xi_{0f}, \theta_2, \delta_{jf}, \mu_f, l_{jf}\, k, j = 1, 2, \ldots, n, k \neq j, f = 1, 2, F)$. Denote the set of all brand switches by $\{j, k\}$. Then, the likelihood of Θ for household i, $L_i(\Theta \mid X)$, is given by:

$$L_i(\Theta \mid X) = \left[\prod_{\{j,k\}} \{\lambda_{ijk}(t \mid X, \Theta) S_{ijk}(t \mid X, \Theta)\} \left\{ \prod_{\substack{l=1 \\ l \neq k}}^{n} S_{ijl}(t \mid X, \Theta) \right\} \right] S_{ij}(t \mid X, \Theta), \qquad (3)$$

where S_{ijk} is the survivor such that,

$$S_{ijk}(t \mid X, \Theta) = \exp\left(-\int_0^t \lambda_{ijk}(u) du \right). \qquad (4)$$

And the other survivors are similarly defined. The term in the square bracket in Equation (3) refers to the likelihood of the $j \rightarrow k$ switch. The last survivor in Equation (3) accounts for right censoring of the data. The sample likelihood function is then given as the product of the likelihood functions of the individual panelists.

[3] It is important to note that in the absence of heterogeneity, i.e., if $\mu_i = \mu$ for all i, the decomposition above is not identified as one could arbitrarily rescale the two components to obtain the same product.

[4] In the empirical analysis, we estimated two- and three-dimensional maps. We found that a higher dimensional map revealed no significant improvement in model fit based on the Bayesian Information Criterion.

[5] We thank a reviewer for bringing this to our attention.

Specification of Unobserved Heterogeneity
As noted previously, households are assumed to be heterogeneous in a subvector of the vector of unknown parameters, Θ_i, for each of the three model formulations. In our empirical analysis, the distribution of this subvector of Θ_i across households is approximated by a discrete distribution with a finite number, S, of supports (Heckman and Singer 1984). The number of supports for the data set we use is determined by increasing S until the value of the Bayesian Information Criterion (BIC) declined when going from S to $S + 1$. Given the "optimal" number of segments, S, the next task would be to obtain market structure implications from the model parameters. For this, we examine the estimated locations of the brands in the perceptual map to evaluate how similar or discrepant are the implications from the three different approaches.

Estimation
Recall that Θ_i denotes the set of parameters to be estimated for household i. Now, Θ_i for household i is a random draw from a discrete distribution with a finite number of supports, S. This requires us to estimate the supports of the distribution, Θ_s, $s = 1, 2, . . , S$ and the associated probability masses, p_s, $s = 1, 2, \ldots, S - 1$ (as the probabilities sum to one). The number of supports, S, that provides the best fit to the empirical data is then determined using BIC. The sample likelihood function will be given by:

$$L(\Theta \mid X) = \prod_{i=1}^{M} \left\{ \sum_{s=1}^{S} L_i(\Theta_i \mid X) p_s \right\}, \tag{5}$$

where M is the sample size. The parameters corresponding to each segment can be estimated using maximum likelihood methods. The EM algorithm can also be used for the purpose. Note that the parameters need to be estimated with several different starting values to ensure that the likelihood is indeed maximized.

As noted in Chintagunta (1994), two approaches have typically been adopted to determine the "optimal" number of dimensions (F): (i) based on interpretability and (ii) based on model fit. The former approach typically leads to focusing on two-dimensional maps. In our empirical applications, we find that using the model fit criterion (BIC) resulted in the two-dimensional model having a lower BIC value than a three-dimensional model. Hence, the recommendations based on both approaches lead to two-dimensional maps being considered.

Parameter Identification
Given observations of repeat purchase times (i.e., when $j = k$ and $\alpha_{ijk} = 0$) of the n brands in the sample, the repurchase parameters, δ_{jf} in the hazard functions are all identified. Turning to the "switch" parameters, α_{ijk}, note that for the no heterogeneity case (i.e., when $\alpha_{ijk} = \alpha_{jk}$ for all i), there are a total of $n(n - 1)$ parameters. These $n(n - 1)$ parameters are identified if we have observations on all possible $j \rightarrow k$ switches. If heterogeneity is accounted for, then S supports will lead to $n(n - 1)S$ parameters (as there will be $n(n - 1)$ parameters corresponding to each support). This will be the number of parameters that have to be estimated under the original Vilcassim and Jain (1991) specification.

In the proposed approach, what we have done is to impose a factor structure on the distribution of the α_{jk} parameters across the households (see Elrod and Keane 1995). For example, if the number of factors $=$ F, then the number of estimated distance parameters $(d_{jk}) = Fn(n - 1)$ and the number of inertia/variety-seeking parameters $= FS$. The parameters will therefore, be identified if

$$Fn(n - 1) + FS \leq n(n - 1)S$$

or when

$$F \leq S \ / \ (1 + (S/(n(n - 1)))). \tag{6}$$

This implies that the number of factors is limited by the extent of heterogeneity in the data. In the limiting case with a large number of brands, it implies that the number of supports provides an upper bound to the number of factors. If (6) is satisfied, then imposing the factor structure results in a more parsimonious specification as compared to the VJ model. When $F = 2$, the total number of estimated parameters are $2n(n - 1) +$ $2S$. Our proposed specification in two dimensions

$$\xi_{1ijk} = \delta_{ik} + \mu_{i1}(l_{j1} - l_{k1})^2 + \mu_{i2}(l_{j2} - l_{k2})^2$$

further simplifies the model by computing the distances based on relative locations of brands in two-dimensional space. Such a decomposition has also been used in a number of previous studies including

Chintagunta (1994), Elrod and Keane (1995), and Erdem (1996). The number of estimated parameters are the following:

Location parameters for the n brands in two dimensions: $2n - 3$

Inertia/variety-seeking parameters for S supports: $2S$

Total number of parameters: $2(n + S) - 3$

There are $2n$ location parameters (n brands in two dimensions). One brand is constrained to lie at the origin (translational invariance). Another brand is constrained to lie along one of the map's axes (rotational invariance). This reduces the number of parameters by three. Hence, there are $2n - 3$ estimated location parameters. The scale invariance property of the map is addressed as the inertia/variety-seeking parameters are constrained to lie between -1 and $+1$. Hence, for identification, we require the following condition to hold: $2(n + S) - 3 \leq n(n - 1)S$. It is straightforward to verify that this condition is always satisfied in most practical applications, i.e., when number of brands exceeds three and there are at least two support points. The above condition also demonstrates how our proposed modification of the Vilcassim and Jain model is more parsimonious. For example, with eight brands and six supports, we need estimate only 25 parameters. The Vilcassim and Jain model by contrast requires the estimation of 336 parameters. It is also easy to verify that $2(n + S) - 3 < 2n(n - 1) + 2S$ for all feasible n. In other words, our proposed specification (2′) is more parsimonious than the factor decomposition in (2) when $F = 2$.

In terms of being able to separately identify the μ (inertia/variety-seeking) and l (location) parameters, note that identification is facilitated by constraining the location parameters (l) to be *household invariant* while allowing households to vary in their inertia and variety-seeking tendencies (i.e., the μ parameters).

Brand-Choice Probabilities

Suppose at time $(t - 1)$, the household purchased brand j. Then the hazard of entry into the product category at time t, conditional on brand j purchase at time $(t - 1)$, is given by the sum of the hazards of entering each of the states, $l = 1, 2, \ldots, n$, from the origin state j ($\Sigma_{l=1}^{n} \lambda_{ijl}(t)$). Kalbfleisch and Prentice (1980) provide the switching probability of purchasing brand k on occasion t conditional on j being purchased at $(t - 1)$ as follows (see their Equation (7.12)):[6]

$$P_{ijk}(t) = \frac{\lambda_{ijk}(t)}{\sum_{l=1}^{n} \lambda_{ijl}(t)} = \frac{(\gamma_{1ijk}^{-\gamma_{0i}} + t^{\gamma_{0i}})^{-1}}{\sum_{l=1}^{n} (\gamma_{1ijl}^{-\gamma_{0i}} + t^{\gamma_{0i}})^{-1}}, \quad (7)$$

where γ_{1ijk} in Equation (7) is given by:

$$\gamma_{1ijk} = \exp(\delta_{ik} + \sum_{f=1}^{F} \mu_{if}(l_{jf} - l_{kf})^2 + X_{ijk}\theta_2). \quad (8)$$

Note that in a traditional brand-choice model, the brand-choice probabilities remain fixed as long as preferences and marketing variables do not change. The above expression implies, however, that brand-choice probabilities are changing over time and there is some underlying temporal pattern to these probabilities (based on the distribution of t). One way of interpreting this is that households have a dynamic component of preferences that affect their choices over time. Allowing for this dynamic in the brand-choice probabilities could affect the implications obtained from the model. The above purchase probability expression thus provides the theoretical linkage between purchase timing and brand choice within the context of our brand-purchase timing hazard model. In the appendix, we discuss an alternative formulation of purchase timing hazards. That formulation leads to a decoupling of purchase timing and brand choice when one focuses on the latter aspect of purchase behavior, resulting in the standard logit model. All the empirical results presented in the next section are also available for such a formulation and can be obtained on request.

According to the proposed model, the switching probability also depends upon γ_{0i}. In Equation (8), heterogeneity across households in their preferences and in their responses to marketing activities is accounted for via the formulation for γ_{1ijk} (as with logit brand-choice models). The term γ_{0i} on the other hand captures the effects of heterogeneity in households' purchase

[6]And no purchase between $(t - 1)$ and t.

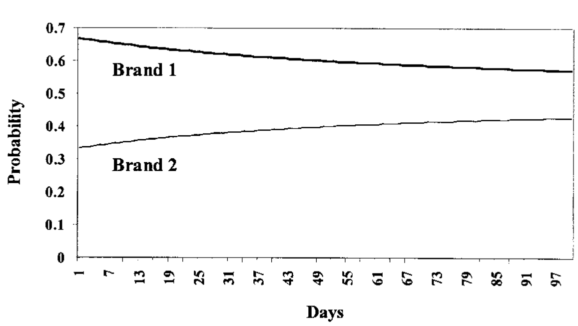
timing behaviors on brand-choice probabilities. In other words, at the same point in time (i.e., t is the same) two households with the same preferences and sensitivities to marketing variables can have different purchase probabilities by virtue of the heterogeneity in their γ_{0i} parameters. This additional source of heterogeneity is not accounted for in traditional brand-choice models.

We now study the implications of the above specification for the (conditional on time) brand-choice probabilities. Consider a two-brand market. As we want to focus on the temporal aspect, assume that marketing and other variables are held fixed over time. For a given inertial household, assume that the previous brand purchased is brand 1 (a similar analysis can be performed for a variety-seeking household). So, $\gamma_{1i11} > \gamma_{1i12}$. Let $\gamma_{1i11} = 2$ and $\gamma_{1i12} = 1$. Under the standard logit assumption, $P_{i11} = 2/3$ and $P_{i12} = 1/3$. When $t = 0$ and $\gamma_{0i} = 1$, Equation (7) yields identical probabilities as the logit model. However, as t increases, households become increasingly indifferent between the two brands, and when $t \to \infty$, $P_{i11} = P_{i12} = \frac{1}{2}$. For the normal average interpurchase time, however, this reduction in preference differential is small as Figure 1 indicates.

It is also straightforward to verify the effects of changes in γ_{0i} on the brand-choice probabilities. As the value increases beyond 1, the initial probabilities (at time $t = 0$) move away from each other (toward 1 and 0, respectively). With a reduction in its value below 1, however, the initial probabilities move toward one another and closer to $\frac{1}{2}$ each.

The brand-choice probabilities unconditional of the interpurchase times can be written as follows:

Figure 1 **Brand-Choice Probabilities Over Time**



$$P_{ijk} = \int_0^\infty \frac{(\gamma_{1ijk}^{-\gamma_{0i}} + t^{\gamma_{0i}})^{-1}}{\sum_{l=1}^{n} (\gamma_{1ijl}^{-\gamma_{0i}} + t^{\gamma_{0i}})^{-1}} f(t)dt,$$

where $f(t)$ is the density corresponding to the distribution of t. It is clear from the above expression that the probabilities do not suffer from the IIA property. This can be verified by writing out the ratio of switching from brand j to brands k and l. This ratio will depend upon the characteristics of the other brands in the choice set. Again, this is in contrast with the probabilities obtained from the proportional hazard model described in the appendix.

3. Empirical Illustration

We use three different scanner panel data sets to fit the purchase timing model described in the previous section. These are (i) the laundry detergent data from the Sioux Falls, SD, market during the period from 1986 to 1988; (ii) the laundry detergent data from the Springfield, MO, market during the period from 1986 to 1988; and (iii) the nondiet soft drinks data from a national panel. For purposes of the empirical analysis, we focus on the four leading (market share by volume) brand sizes of liquid and powdered detergents in each of the two markets; and the top seven brand sizes of soft drinks in that panel. The included brands accounted for over 68% of purchases in all three categories. The data are for 4,377 purchases made by 807 panelists for the Sioux Falls data; 8,058 purchases by 810 panelists for the Springfield data; and 6,765 purchases made by 827 households for the soft drinks category. The descriptive statistics for these data are in Table 1. The numbers for feature and display denote the proportion of purchases made on these promotions.

From Table 1, we see that only two of the top four brands of liquid laundry detergents are common across the two geographic markets (Tide 64 and Wisk 64). Further, among the powdered detergents, the only common brand size is Tide 42.[7] Also, we are not at

[7]Note that while this may not allow us to establish the commonality (or lack thereof) of the brands' perceptual locations across the two markets, it nevertheless allows us to make category-level comparisons. For example, is the nature of inertia and variety seeking similar across markets for the same category, is detergent (category) purchase timing behavior comparable, etc.?

Table 1 Descriptive Statistics for the Data

Brand (Liquid)	Price	Feature	Display	Share	Brand (Powder)	Price	Feature	Display	Share
				Detergents (Sioux Falls)					
Wisk 64	5.108	0.179	0.085	0.124	Tide 72	4.676	0.168	0.088	0.181
Surf 64	5.379	0.142	0.079	0.146	Tide 42	5.018	0.032	0.027	0.145
Tide 64	6.067	0.123	0.061	0.119	Tide 147	4.534	0.049	0.037	0.091
Tide 96	5.888	0.178	0.015	0.091	Oxydol 72	4.992	0.118	0.031	0.104
				Detergents (Springfield)					
Wisk 64	5.437	0.027	0.018	0.060	Tide 42	4.644	0.257	0.370	0.480
Wisk 32	6.109	0.010	0.013	0.047	Surf 42	4.842	0.064	0.109	0.152
Era Plus 64	5.885	0.011	0.000	0.042	Cheer 42	4.741	0.038	0.096	0.091
Tide 64	6.107	0.010	0.082	0.041	Purex 42	2.859	0.054	0.196	0.087
				Soft Drinks					
Brand 1	17.853	0.077	0.452	0.258	Brand 32	4.692	0.113	0.010	0.123
Brand 21	13.464	0.057	0.545	0.199	Brand 4	27.457	0.031	0.184	0.089
Brand 22	11.289	0.267	0.283	0.071	Brand 5	29.001	0.073	0.261	0.076
Brand 31	3.738	0.305	0.088	0.183					

Notes. The numbers for feature and display denote the proportion of purchases made on these promotions. Prices for the detergent data are per ounce; prices for the soft drinks data are per 6 ounces. Detergent brands in the first column are liquid; the others are powdered.

liberty to reveal the identities of the soda brands. The brands are numbered, however, to identify variants of the same brand, e.g., brands 21 and 22 are two variants of brand 2; and brands 31 and 32 are two variants of brand 3. Further, brand 3 is the only private label brand.

In the discussion of the empirical results that follow, we first present some diagnostic tests we carried out, then the statistics on model fit, followed by an analysis of the dynamic aspects of purchase behavior, i.e., inertia and variety seeking. Next, we discuss the purchase timing behavior of households. Then we provide results on the market structure implications derived from the model estimates. This is followed by predictive validation. Subsequently, we discuss the results from an application of the methodology to predicting the perceptual location of a line extension. Then, we obtain implications on the timing of promotions obtained from the proposed model. We conclude this section with a comparison of the market structure implications obtained when only brand-choice information is used.

Diagnostics

Several diagnostics were performed. First, we tested for the sensitivity of the estimated results to the number of factors included in the specification of heterogeneity (i.e., F in the previous section). In all cases, we found that $F = 2$ fit the data the best. In other words, including more than two factors did not improve the model fit. It is important to note, from the previous section, that the "optimal" number of supports (S) also depends on the number of factors included in the model (for identification purposes).

In addition, we tried to ascertain whether the proposed parsimonious model, after adjusting for the estimated number of parameters, fit the data as well as the full VJ model. It is important to note here that under the VJ specification, one needs to observe a "sufficient" number of switches from A to B in order to estimate the "A to B" hazard for that pair (similarly for the B to A hazard). This is not required of the proposed model specification. In Table 2, we provide the brand-switching matrix for the soft drinks product category. It is clear from this matrix that it would be extremely

Table 2 Brand Switching Matrix: Nondiet Soft Drinks

	Brand 1	Brand 21	Brand 22	Brand 31	Brand 32	Brand 4	Brand 5
Brand 1	848	272	31	89	46	140	111
Brand 21	197	585	71	86	48	128	47
Brand 22	45	74	249	20	29	9	5
Brand 31	123	79	17	795	72	28	25
Brand 32	59	49	38	65	497	13	13
Brand 4	127	58	10	17	10	190	89
Brand 5	138	56	4	22	9	34	172

difficult to estimate a unique set of parameters for the brand 5 → brand 22 transition, given the paucity of observations on that switch. Even ignoring the transitions for which we had fewer than 10 observations, we find that the proposed model, after adjusting for the number of estimated parameters, fits the data better than the VJ specification (i.e., using BIC). Similar results hold for the other two data sets also.

Discussion of Model Fit

Table 3 provides the values of the BIC for an increasing number of supports for the heterogeneity distributions for each of the three data sets. Table 3 reveals that the BIC drops steadily in going from the no heterogeneity model (one support distribution) to a model with a higher number of supports for each of the three model formulations across the three data sets. The BIC is minimized at four supports for the Sioux Falls data; three supports for the Springfield data; and four supports for the soft drinks data. This implies, consistent with the previous literature (Chintagunta 1994), that a multiple segment solution is required to adequately capture the heterogeneity in these data sets. The implications of the heterogeneity distributions are discussed later.

Analysis of Inertia and Variety-Seeking Behavior

As described in the model formulation section, the model parameters provide insights into the nature of inertia and variety-seeking tendencies of households. Such implications are obtained via the interpretation of the μ_{sf} parameters (segment s, attribute f). A positive value implies variety-seeking behavior, whereas a negative value corresponds to inertia. First, we discuss the overall results obtained from the three data sets and then discuss the segment level results. Table 4 provides the mean values for the distribution of inertia/variety-seeking effects across the three sources of data.

Table 3 Model Fit Statistics: BIC

Number of Supports	Detergents (Sioux Falls)	Detergents (Springfield)	Nondiet Soft Drinks
1	8,001	16,212	14,977
2	7,323	14,429	14,174
3	7,315	14,069*	13,998
4	7,268*	14,134	13,879*
5	7,293		13,944

Note. BIC = −2*Log-Likelihood + (# Parameters)*ln(# Observations).

*Models for which BIC is minimized. The number of parameters for these models are, respectively, 51, 43, and 44.

Table 4 Mean Values for the Distribution of Inertia/Variety-Seeking Effects

	Mean Value	
Data Set	Attribute 1	Attribute 2
Detergent—Sioux Falls	−0.129	−0.238
Detergent—Springfield	−0.802	0.920
Soft drinks	0.303	−0.398

First, we note that there appears to be considerable heterogeneity in the nature of purchase dynamics across the three market categories. We find that households in Sioux Falls are "mildly" inertial in the two attributes that drive their purchases. On the other hand, for the same detergent market, households in Springfield behave quite differently. In other words, they are, on average, strongly inertial in one attribute and strongly variety seeking in the other. At least two possible explanations exist. One is that households are very different across the two markets in the way they behave with respect to the same two attributes. The other is that the attributes that determine household preferences are very different across the two markets. A limitation of our analysis is that we cannot separate out these two explanations. For the soft drinks data, based on a national sample, we find that consumers display moderate amounts of inertia in one attribute and some variety seeking in the other attribute. This finding is quite consistent with our prior expectations of the soft drinks category.

In Table 5, we provide details of the estimated inertia/variety-seeking parameters. Table 5 indicates that there is considerable heterogeneity in the extent of inertia and variety seeking across supports for two of the three data sets. Specifically, we find this to be the case in the detergents (Sioux Falls) and soft drinks product markets. In the former data, we have households that exhibit zero-order behavior (support 2), strong inertia (support 3), and mild inertia (support 4). Furthermore, support 1 households exhibit zero-order behavior along one attribute and inertia along the other. Turning to the soft drinks data, we see that support 4 households show inertial behavior along both attributes. However, households associated with the other three supports show varying levels of inertia along attribute 2 and different levels of variety seeking along attribute 1.

The households that show the least amount of heterogeneity are those from the Springfield market. To obtain some further insights into the factors driving the three-support solution in this case, we examined the distribution of marketing variable effects across these three supports. The findings provide us with some rationale for the heterogeneity in these data. Our estimates indicate that households associated with support 1 are the most sensitive to marketing activities (price and feature), whereas those associated with support 2 are the least. Households in support 3 exhibit an intermediate amount of heterogeneity along these two marketing variables. In all cases, we found the display effect not to be significant. To summarize, our model formulation does appear to be able to distinguish between the two kinds of purchase dynamics—inertia and variety seeking.

Purchase Timing

As noted previously, one of the key advantages to the model specification is that it allows for different transition or hazard rates across different brand pairs. This is a property that the formulation shares with Vilcassim and Jain (1991), although with fewer parameters. To demonstrate the ability of the model to allow for different rates across brand pairs, we plot the transition rates in Figures 2, 3, 4, and 5. We focus first on the detergents data set from Sioux Falls. In the interests of space, we do not show the transitions from in and out of each of the eight brands. We focus on the largest powdered detergent brand (Tide 72) and the largest liquid detergent brand (Wisk 64) and examine the rates of exit from these two brands. Note also that we have allowed for heterogeneity in transition rates across supports. For purposes of illustration, we present the rates corresponding to a support that is characterized by a high level of inertia (high inertia support—Figures 2 and 3) and a support that is characterized by a low inertia level, i.e., close to zero-order behavior (low inertia support—Figures 4 and 5). According to Equation (1), the hazard functions depend upon the covariates price, display, and feature. One option is to set these variables at their average values in plotting the hazards. However, this will not enable us to compare the intrinsic switching rates sans marketing variables. Consequently, the hazard rates in the figures are drawn after setting all the marketing variables to zero.

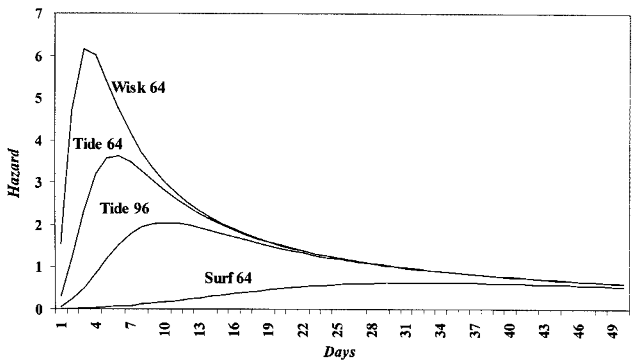
Figure 2 for Wisk 64 and Figure 3 for Tide 72 in the

Table 5 **Parameter Estimates and (Standard Errors) for Distribution of Inertia/Variety Seeking**

	Detergents (Sioux Falls)			Detergents (Springfield)			Soft Drinks		
Support	Attribute 1	Attribute 2	Weight	Attribute 1	Attribute 2	Weight	Attribute 1	Attribute 2	Weight
1	0.027*	−0.141	0.283	−0.868	0.920	0.567	0.328	−0.657	0.058
2	−0.059*	0.081*	0.068	−0.701	0.758	0.063	0.838	−0.192	0.360
3	−0.992	−0.967	0.092	−0.720	0.949	0.370	0.157	−0.137	0.333
4	−0.074	−0.206	0.557				−0.282	−0.985	0.249

*Not significant at the 5% level of significance.

Figure 2 **Transition Rates from Wisk 64 to Other Liquids (High Inertia Support)**



Transition Rates from Wisk 64 to Powders (High Inertia Support)

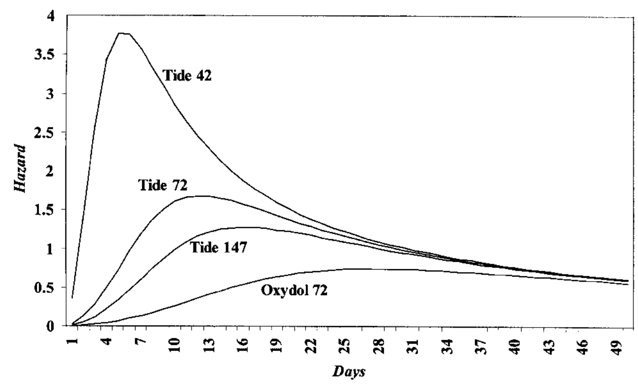
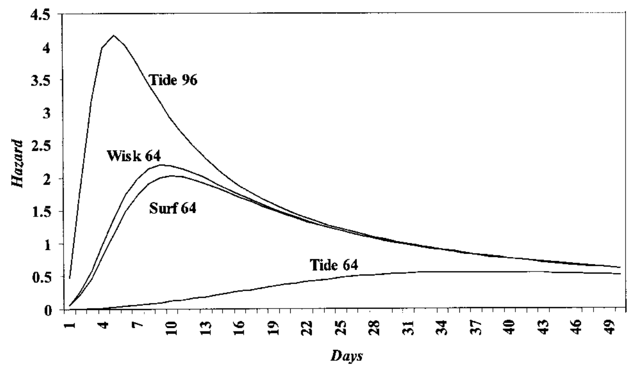
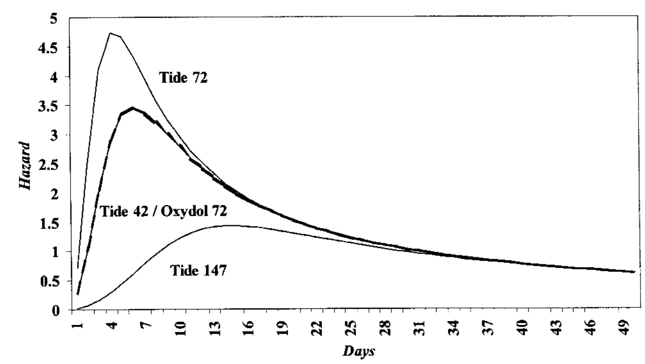


Figure 3 **Transition Rates from Tide 72 to Liquids (High Inertia Support)**



Transition Rates from Tide 72 to Other Powders (High Inertia Support)

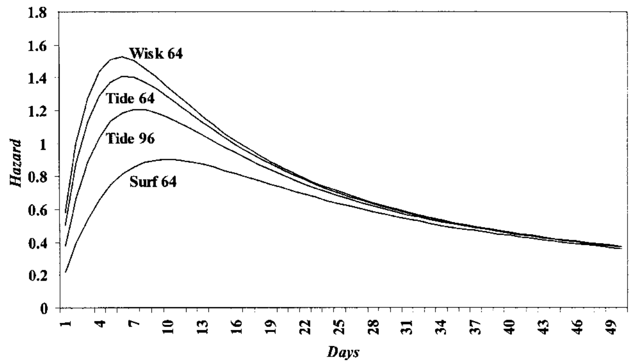


high inertia setting reveal the following: Consistent with previous literature (Jain and Vilcassim 1991; Vilcassim and Jain 1991), all the transition rates exhibit a nonmonotonic or a monotonically increasing pattern. The nonmonotonicity has been interpreted by previous authors as suggesting regularity in households' purchases of brands. In addition, we find that the switching rates vary across destination states, i.e., brand forms. For example, in the case of Wisk 64, the most likely outcome is a repeat purchase of the same brand, followed by Tide 64 (a liquid) and Tide 42 (a powder) and Tide 96 (a liquid). The high inertia setting implies that repeat purchase is most likely. Further, Wisk 64 does not appear to be competing much with Lever's other liquid brand, Surf 64. This finding makes sense for Lever's perspective. On the other hand, it also indicates competition from a number of rival P&G's brands—specifically, Tide in the same size pack.

Similarly, from Figure 3, we see that a purchase of Tide 72 (a powder) is most likely followed by a purchase of the same brand. Once again, this is consistent with the high inertia condition. The next most likely brand size to be purchased appears to be Tide 96—a liquid. From P&G's category management perspective, this indicates that the firm is not losing sales to rival Lever brands. However, there is the issue of cannibalization. The next most likely brands are, once again, P&G brands Oxydol 72 and Tide 42—both powders.

In the low inertia state, we would expect the transition rates to reflect the relative sizes of the destination states, after controlling for the effects of marketing variables. Consistent with this expectation, we find from Figures 4 and 5 that Wisk 64 is the most likely destination state, whether or not the origin state is Wisk 64 (a liquid) or Tide 72 (a powder). This is what one would expect from (close to) zero-order purchase

Figure 4 Transition Rates from Wisk 64 to Other Liquids (Low Inertia Support)



Transition Rates from Wisk 64 to Powders (Low Inertia Support)

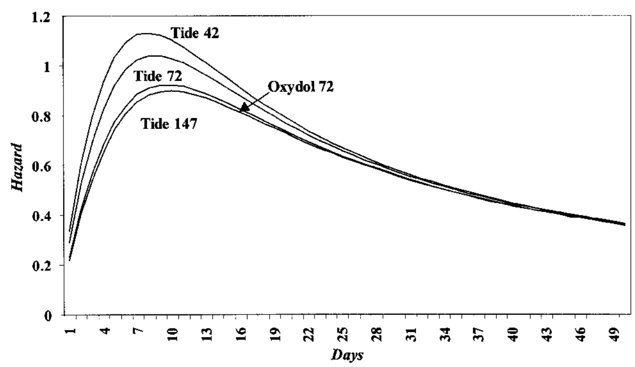
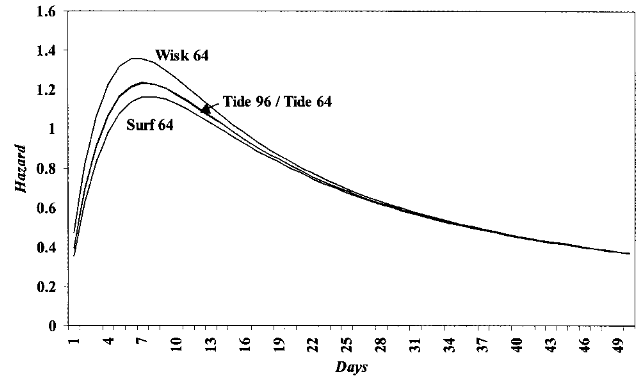
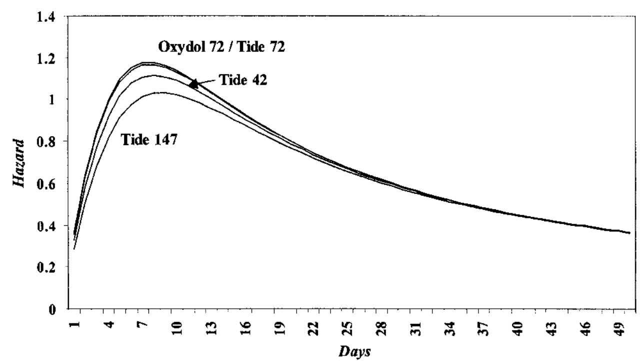


Figure 5 Transition Rates from Tide 72 to Liquids (Low Inertia Support)



Transition Rates from Tide 72 to Other Powders (Low Inertia Support)

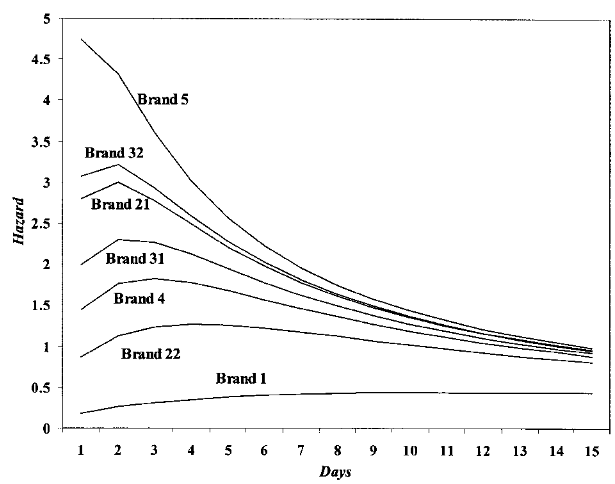


behavior. In both figures we find that the next two most likely brands are Tide 64 and Tide 96. Of course, the actual transition rates do vary as the inertia parameters are -0.07 and -0.21 for attributes 1 and 2, respectively.

In the interests of space, we do not provide the transition rates from the other two data sets. However, it is of some interest to contrast the transition rates from inertial households to those from variety-seeking households. Accordingly in Figure 6, we provide the transition rates from brand 1 in the soft drinks category. The figure corresponds to households in support 2 who are characterized by variety-seeking behavior along one attribute and inertia along the other (see Table 5).

Figure 6 illustrates the nature of variety-seeking behavior. Households that are variety seekers are now much less inclined to purchase brand 1 after a brand 1

Figure 6 Transition Rates from Brand 1 (Variety Seeking Support)

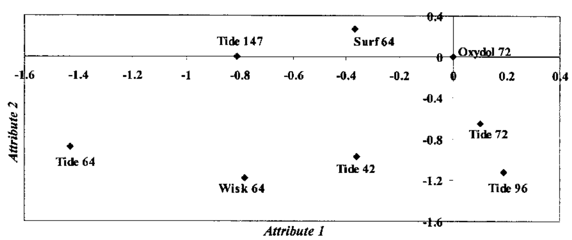


purchase. Rather, all six other brands are more likely to be purchased. From the results in this subsection, we find that differences across households translate into differences in their purchase rates for the different brands, with the nature and extent of inertia and variety seeking affecting these purchase rates.

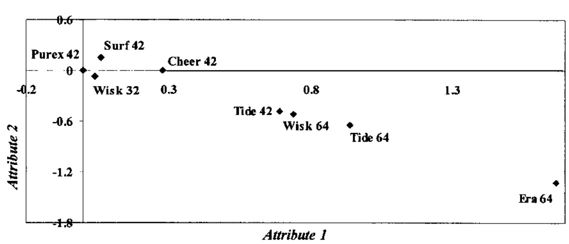
Market Structure Implications

In Figure 7, we provide the two-dimensional maps estimated for each of the three data sets. It is important to note that the attributes of the product spaces need arbitrary nomenclature. Regarding the two detergent data sets, we find that the method delineates the liquid laundry detergents from the powdered ones. Specifically, for the Sioux Falls data we find that the two forms are separated along the vertical axis (with the exception of Surf 64, a liquid, and Tide 42, a powder).

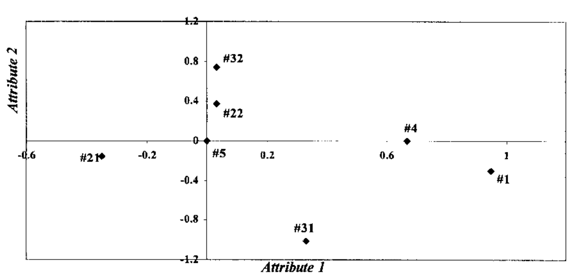
Figure 7 **Detergents (Sioux Falls)**



Detergents (Springfield)



Soft Drinks



Similarly, for the Springfield data, the separation occurs along the horizontal axis (with the exception of Wisk 32). Furthermore, for the Springfield data, the brands appear to be delineated along the size dimension also, unlike the Sioux Falls data.

The two-dimensional map for Sioux Falls provides some interesting insights when viewed from the perspective of the Lever brands. It appears that the two Lever brands, Wisk and Surf, are positioned to compete against different P&G brand form sizes. While Wisk is seen as competing with Tide 64 (a liquid) and Tide 42 (a powder) the most, Surf 64 is perceived as being similar to Tide 147 and Oxydol 72, both P&G powdered brands. Comparing these findings with those from the Springfield data, we find that, once again, Wisk 64 is seen as competing most closely with Tide 64 and Tide 42. This is one key point of similarity between the two detergent markets. A significant point of difference is the competition among the smaller detergent sizes. It appears that Purex 42, Surf 42, and Cheer 42 (all powders) are competing with one another and with Wisk 32 (a liquid).

The map based on the soft drinks data is more difficult to interpret as we are not at liberty to divulge brand names. However, it does appear as if brands 5, 22, and 32 are locked in competition, whereas 1, 4, 21, and 31 seem much more insulated from their rivals. Regarding the perceived distance between 31 and 32, we note the attribute along which they differ the most is attribute 2. So while one variant of the private label brand (32) competes with two national brand variants (22 and 5), the other (31) is not perceived as being a threat to the national brands.

Predicting the Perceptual Location of a Line Extension

This exercise is carried out using the Sioux Falls data. Recall that the data set has eight brand sizes. These come from four brands (Wisk, Surf, Tide, and Oxydol), two forms (powder and liquid), and five sizes (64, 96, 42, 72, and 147 oz.). These three, in effect, represent only two independent features as the information on product form is fully contained in information on the size. For example, knowing that size is 72 oz. reveals that this is a powdered detergent. In this case, the different brands and sizes together yield 20 possible brand-size combinations. Of these, only eight are included in our data set. This implies that there is an opportunity to predict the perceptual locations and shares of the 12 remaining brand-size combinations.

An important caveat here is that a likely reason for the firms not launching these 12 brand sizes is a low demand assessment based on internal marketing research. Furthermore, the prediction methodology used here does not take into account interaction effects between brand name and size, which could influence preferences for brand-size combinations. For example, buyers of brand X may be small households who prefer small package sizes. In which case, it is unlikely that the firm will introduce larger size packages. Hence, our predictions here need to be construed as only a starting point based on the demand for the existing brands. This prediction takes into account the similarity/dissimilarity of the new brand to the existing ones.

The perceptual location parameters (l_{jf}) and the brand-specific parameters (δ_j) are decomposed as follows: for the δ_j parameters: with four brands and five sizes, we have three unique brand parameters (with one normalized to zero) and four unique size parameters (with one normalized to zero); we have a total of seven parameters (identical to the number of unique parameters with eight brand-sizes estimated previously). Similarly for the location parameters, there is a total of 13 unique parameters, as with the model already estimated. Hence, the Fader and Hardie decomposition yields exactly as many parameters as without the decomposition.

The estimation results are as follows: A four-support distribution was found to be adequate for this model also. The fit of this model (BIC = 7,289) is slightly poorer than the BIC value in Table 3 (7,268). To verify whether the market structure implications are consistent with those obtained without the decomposition, we correlated the interbrand distances obtained with those from Figure 7. The correlation was 0.97, implying extremely high consistency. In Figure 8, we plot the perceptual locations of the eight brand sizes obtained from the Fader and Hardie decompositional approach. As in Figure 7, we find a demarcation between powdered and liquid detergents (with the exceptions of Surf 64 and Tide 42 as before).

Figure 8 also shows the predicted location of a possible "line extension" by Procter & Gamble. This is the liquid Oxydol 64 brand size. We find that the predicted location is closest to Tide 64 and Tide 147, both P&G brands. The nearest rival brand is Lever's liquid Surf 64 brand size. Hence, it could potentially draw share away from this brand. The tradeoff, of course, is in terms of the cannibalization that would result for the other P&G brand form sizes. We also see that although the current Oxydol is a powdered detergent, performing the decomposition into brand and size, provides a prediction for liquid Oxydol as being perceptually similar to other liquid detergents, while at the same time, sharing similarities with the Oxydol brand name.

We could also perform a competitive draw/cannibalization computation to aid in making the tradeoff. However, this would be conditional on a specified interpurchase time (recall the brand-purchase probability expressions derived earlier). Based on Figure 8, one would expect the shares of Oxydol 64 to come from brands located close to its location, all else being equal (i.e., marketing variables price, display, and feature). However, note that it would also depend upon the extent of variety-seeking behavior, as these households would switch from brands located far away in perceptual space. This illustrates that the Fader and Hardie approach can be extended to market structure implications in the context of purchase timing models.

Implications for Timing of Promotions

In this section, we take the perspective of a direct marketer who has access to the brand choices and purchase times from a sample of households in the market. In other words, while the information contained in the data are similar to those we use in the estimation above, the context is within a category of interest to a direct marketer. In such a case, the direct marketer knows (a) the brand purchased, (b) the timing of the purchases, and (c) based on the model estimation, whether the customer under consideration is inertial or variety seeking. One of the implications from the proposed model is that the probability of brand switching varies over time. Hence, the direct marketer could exploit this information to time promotions and promotional material appropriately. Consider an inertial household in this market. The brand-switching

Figure 8 Predicted Location of Line Extensions

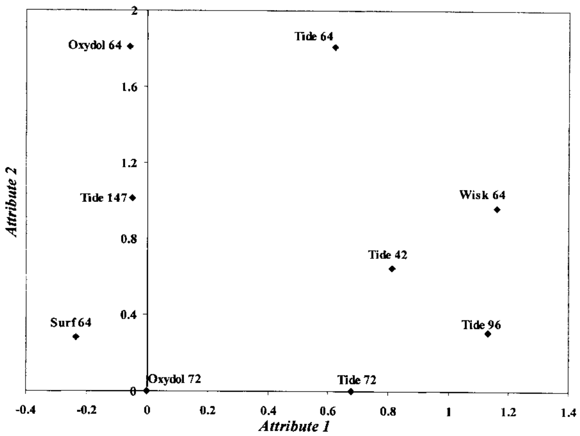
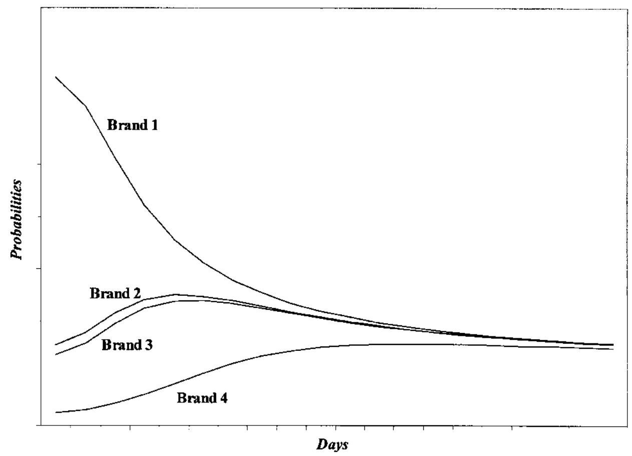


Figure 9 Transition Probabilities from Brand 1 (High Inertia)



probabilities in a four-brand category could look like those in Figure 9. If the household purchases brand 1, then due to inertia, the likelihood of purchasing brand 1 on the next purchase occasion is the highest. A model based only on brand-choice data would also give us a similar result. In addition, the proposed model implies that the probability of the brand 1 to brand 1 "switch" varies over time for that household. This variation is depicted in Figure 9. The transition probabilities in Figure 9 would be based on Equation (7).

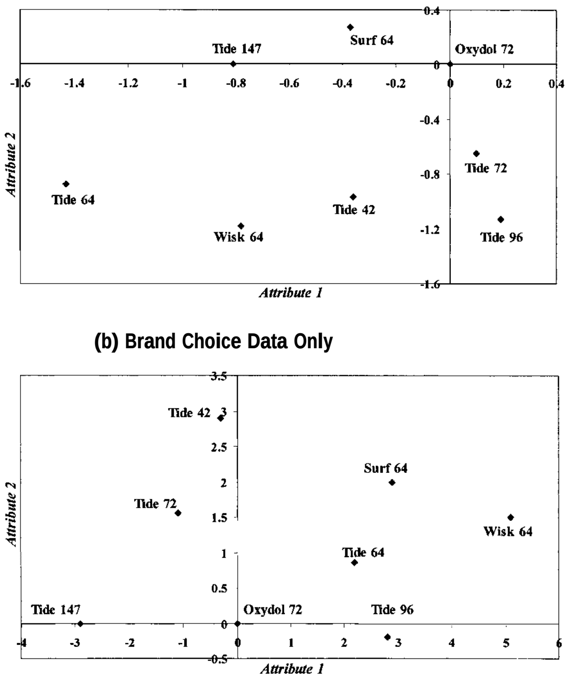
Figure 9 indicates that the brand 1 to brand 1 switching probability diminishes over time for the inertial household. And over time, the probabilities of purchasing the various brands converge toward one another. This has implications for the timing of promotions by competing brands. Figure 9 shows that right after a brand 1 purchase, the probability of buying brand 1 is the highest. Furthermore, the difference in brand 1 and brand 2 probabilities is quite high. So the it would not be optimal to promote brand 2 when the probability difference is so high because the promotion effect may not be large enough to bridge the gap in the probabilities. However, after a while, the probability of brand 1 has declined to such an extent that the difference in probabilities is "small enough." At this stage, a promotion by brand 2 raises its transition probability higher than that for brand 1. Hence, a switch can be induced. One has to make a tradeoff in terms of how long one should wait to promote brand 2. It may not be in the brand's best interests to wait too long as then the probabilities of the other brands would "catch up," making it more difficult to effect a defection to brand 2. Hence, the model results provide implications for the "optimal" timing of promotions by the competing brands that cannot be obtained from a brand-choice model without the purchase timing aspect. A similar analysis can be performed for a variety-seeking household.

Comparison with a Model of Brand Choice

Recall from Equation (7) that the (conditional) brand-choice probabilities from the proposed model are explicit functions of the interpurchase times. When purchase times are not accounted for, the probabilities are given in Equation (A.3) of the appendix, i.e., similar to regular logit brand-choice probabilities. In Figure 10, we provide a comparison of market structure implications obtained from the two approaches using the data and results from the Sioux Falls market.

Prima facie, there appear to be differences in the perceptual brand locations obtained from the two approaches. To summarize these differences, we computed the coefficient of congruence between the maps in Figures 10(a) and 10(b). Not surprisingly, the value obtained was 0.28. This implies that the implications are affected by inclusion of purchase timing data. As noted previously, the model underlying Figure 10(a) allows brand-switching probabilities to change over time, with the pattern of variation being brand-pair

Figure 10 **(a) Purchase Timing and Brand Choice Data**



specific. On the other hand, the model underlying Figure 10(b) assumes that these probabilities are time invariant (after controlling for the effects of time-varying marketing variables).

The reason for the difference in market structure implications from the two sets of data can be identified by drawing vertical lines at different time points in Figure 9, with one such line passing through the mean interpurchase time for the household. Then, one can think of the implications coming from the brand-choice data as being driven by the probabilities along the mean interpurchase time line. On the other hand, when using purchase timing data, these probabilities are allowed to differ from those on the mean interpurchase time line in a brand-specific manner. The differences in the brand preferences over time that drive the temporal variation in probabilities account for the differences in market structure implications.

We also obtained similar implications from the other two product markets. In the interests of space, we do not provide the figures for those markets. However, the summary congruence measures for these markets are 0.35 (detergents—Springfield) and 0.46 (soda). While these are higher than the congruence measure for the Sioux Falls market, they continue to indicate large differences in the implications obtained with and without purchase timing data.

4. Conclusions

In this paper, we have integrated the notion of inertia/variety seeking with a model of purchase timing and brand switching. The objective has been to provide greater insight into the nature of state dependence in models of purchase timing. Our empirical analysis demonstrates that the model enables us to distinguish between households that are inertial and those that are variety prone. In addition to accounting for state dependence, the model also accounts for the effects of unobserved heterogeneity among households in their brand preferences and in their sensitivities to marketing activities. We provide some insights into the timing of promotions that can be obtained from our proposed model and obtain implications for market structure.

Integrative frameworks for the analysis of household purchasing behavior have been proposed in the marketing literature (e.g., Chiang 1991). Interestingly, these frameworks have not been fully exploited to reveal greater insights into market structure than are available from the analysis of only brand-choice behavior of households. This study is a first step in the general direction of using additional, relevant, information for the purposes of investigating interbrand competition in the marketplace. While empirical differences were observed in the market structure implications obtained with and without using purchase timing data, we recognize that the theoretical underpinnings of these differences have not been explored in any great detail in this paper. We leave this issue for future research.

There exist other directions for future research that could enhance the generalizability of our results here. Given the availability of other data sets, a further cross-category replication of our results may be useful. Furthermore, investigating alternative specifications for the baseline hazard function would be informative of the robustness of the results to our log-logistic assumption. The choice of a discrete distribution to characterize heterogeneity in the data could restrict the range of heterogeneity captured in the model. A comparison with continuous heterogeneity distributions could be useful. Exploring other substantive implications of the proposed framework could shed further light on the usefulness of the proposed framework.[8]

Appendix

We compare the formulation provided in the text with that from a proportional hazard specification. Under this latter specification, the hazard for household i at time t of switching from brand j to brand k, $\lambda_{ijk}(.)$ can be written as:

$$\lambda_{ijk}(t \mid X, \theta) = \lambda_{0i}(t \mid \theta_{1jk}) \exp\left(\delta_{ik} + \sum_{f=1}^{F} \mu_{if}(l_{jf} - l_{kf})^2 + X_{ijk}\theta_2\right), \quad \text{(A.1)}$$

where $\lambda_{0i}(.)$ is the household-specific, brand-invariant baseline hazard function. It is important to note, however, that $\lambda_{0i}(.)$ does not represent the product category purchase hazard (à la Jain and Vilcassim 1991). That hazard is obtained by summing $\lambda_{0ijk}(.)$ over all origin and destination states. All other terms are as defined previously. The baseline hazard is given by

$$\lambda_{0i}(t) = \frac{\gamma_{0i}\gamma_{1i}^{\gamma_{0i}}\, t^{\gamma_{0i}-1}}{1 + \gamma_{1i}^{\gamma_{0i}}\, t^{\gamma_{0i}}}. \quad \text{(A.2)}$$

Note that the above baseline hazard does not contain the effects of marketing variables and other effects.

As with the proposed specification, the hazard of entry into the product category at time t, conditional on brand j purchase at time $(t - 1)$, is given by the sum of the hazards of entering each of the states, $l = 1, 2, \ldots, n$, from the origin state j $(\Sigma_{l=1}^{n} \lambda_{ijl}(t))$. The probability of purchasing brand k on occasion t conditional on j being purchased at $(t - 1)$ is given by:

$$P_{ijk}(t) = \frac{\lambda_{ijk}(t)}{\sum_{l=1}^{n} \lambda_{ijl}(t)}$$

$$= \frac{\lambda_{0i}(t)\exp\left(\delta_{ik} + \sum_{f=1}^{F} \mu_{if}(l_{jf} - l_{kf})^2 + X_{ijk}\theta_2\right)}{\sum_{l=1}^{n} \lambda_{0i}(t)\exp\left(\delta_{il} + \sum_{f=1}^{F} \mu_{if}(l_{jf} - l_{lf})^2 + X_{ijl}\theta_2\right)}$$

$$= \frac{\exp\left(\delta_{ik} + \sum_{f=1}^{F} \mu_{if}(l_{jf} - l_{kf})^2 + X_{ijk}\theta_2\right)}{\sum_{l=1}^{n} \exp\left(\delta_{il} + \sum_{f=1}^{F} \mu_{if}(l_{jf} - l_{lf})^2 + X_{ijl}\theta_2\right)}. \quad \text{(A.3)}$$

Equation (A.3) is identical to that obtained from the logit brand-choice model. The switching probability depends explicitly on the brand preferences, relative brand proximity (state dependence), and marketing variables, while accounting for inter-household differences. Further, it is similar to the specification in Erdem (1996).

One of the important features to note from the above expression is that the interpurchase times no longer affect the brand-switching probabilities. Our proposed specification, on the other hand, explicitly links the purchase timing and brand-choice decisions in the brand-switching probabilities. Therefore, while we have carried out the empirical analyses with both specifications, we focus on the proposed formulation when discussing the results. Results from the proportional hazard model are available on request.

[8]The author thanks the editor, the area editor, and two anonymous reviewers for numerous helpful comments and suggestions. Acknowledgments to Fusun Gonul for her useful feedback and to Sachin Gupta and PB Seetharaman for the data. The usual disclaimer applies.

References

Chiang, J. 1991. A simultaneous approach to the whether, what and how much to buy questions. *Marketing Sci.* **10** (4) 297–315.

Chintagunta, P. K. 1994. Heterogeneous logit model implications for brand positioning. *J. Marketing Res.* **31** 304–311.

Elrod, T., M. P. Keane. 1995. A factor-analytic probit model for representing market structure in panel data. *J. Marketing Res.* **32** 1–16.

Erdem, T. 1996. A dynamic analysis of market structure based on panel data. *Marketing Sci.* **15** (4) 359–378.

Fader, P. S., B. G. S. Hardie. 1996. Modeling consumer choice among SKUs. *J. Marketing Res.* **33** (November) 442–451.

Fraser, C., J. W. Bradford. 1983. Competitive market structure analysis: Principal partitioning of revealed substitutabilities. *J. Consumer Res.* **10** 15–30.

Grover, R., V. R. Rao. 1988. Inferring competitive market structure based on a model of interpurchase intervals. *Internat. J. Res. Marketing* **5** 55–72.

Heckman, J. J., B. Singer. 1984. A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica* **52** 271–320.

Jain, D. C., N. J. Vilcassim. 1991. Investigating household purchase timing decisions: A conditional hazard function approach. *Marketing Sci.* **10** (1) 1–23.

Kalbfleisch, J. D., R. L. Prentice. 1980. *The Statistical Analysis of Failure Time Data*. Wiley, New York.

McAlister, L. 1982. Variety seeking behavior: An interdisciplinary review. *J. Consumer Res.* **9** 311–322.

——, E. Pesemier. 1982. A dynamic attribute-satiation model for choices made across time. *J. Consumer Res.* **9** (3) 141–150.

Roy, R., P. K. Chintagunta, S. Haldar. 1996. A framework for investigating habits, the hand of the past, and heterogeneity in dynamic brand choice. *Marketing Sci.* **15** (3) 280–299.

Trivedi, M., F. M. Bass, R. C. Rao. 1994. A model of stochastic variety seeking. *Marketing Sci.* **13** (3) 274–297.

Vilcassim, N. J., D. C. Jain. 1991. Modeling purchase-timing and brand-switching behavior incorporating explanatory variables and unobserved heterogeneity. *J. Marketing Res.* **28** 29–41.

This paper was received December 16, 1996, and has been with the author 9 months for 2 revisions; processed by Greg Allenby.