A Composite Transition Model for Brand Choice and Purchase Timing Data

Lynn Kuo and Zhen Chen

Department of Statistics, University of Connecticut, Storrs, CT 06269

(lynn@stat.uconn.edu) and (zhen@stat.uconn.edu)

Abstract

A new approach for analyzing duration state data in brand-choice studies is explored. This approach not only incorporates the correlation among repeated purchases for a subject, it also models the purchase timing and the brand decision jointly. The former is accomplished by applying transition model approaches from longitudinal studies. The latter is done by conditioning on the brand choice variable. Then mixed multinomial logit models and Cox proportional hazards models are employed to model the marginal densities of the brand choice and the conditional densities of the interpurchase time given the brand choice. We will illustrate our findings by applying this new method to Nielsen household scanner panel data.

KEY WORDS: Cox proportional hazards model; Longitudinal analysis; Multinomial logit model; State duration model.

In brand-choice studies, we are interested in factors influencing both the brand-switching patterns and purchase-timing decisions. Usually, panel data are collected for a period of time from N households where we can extract a series of interpurchase times and brands chosen for each purchase of a household. Moreover, we collect a list of covariates at each purchase. The covariates may include characteristics of both the consumers (chooser-specific) and the brands (choice-specific). For example, the choice-specific explanatory variables may include the price of each brand or an indicator of a special display for each brand; the chooser-specific variables may contain household income and size.

Most of the older approaches to analyzing this kind of multiple-spell and multiple-destination data are restricted to marginal approaches where attention is given only to the state-space part (brand choice) of the data or the interpurchase-time part of the data. For example, the former includes the discrete brand-choice models of McFadden (1974) and Jain *et al.* (1994); the latter includes the proportional hazards model (Cox 1972) for the interpurchase time studied by Jain and Vilcassim (1991).

The more advanced studies in this area include (1) the conditional model (for the interpurchase time given the transition states) proposed by Vilcassim and Jain (1991), and Gönül and Srinivasan (1993), and (2) the joint model for the interpurchase time and the brand choice given by Chintagunta and Prasad (1998). Vilcassim and Jain derive a competing risk model where each latent interpurchase time between transitions is modeled by the Cox regression model with frailty. They found that the probability distribution of interpurchase times is not the same for various switching between brands. Gönül and Srinivasan (1993) apply a similar approach except that they restrict the slopes to be the same in the Cox regression model for different transitions. Chintagunta and Prasad apply the dynamic McFadden model (Heckman and Singer 1985), where they model the marginal distribution of the interpurchase time and the conditional distribution of the brand choice given the interpurchase time.

The bivariate model of Chintagunta and Prasad provides further advancement in the area. However, improvement can be made upon their models. First, they fit the same model for the interpurchase time independent of the transitional states. This is unsatisfactory according to the findings of Vilcassim and Jain (1991), because the probability distribution of interpurchase times is usually not the same for various switching between brands. Second, they assume the same model and conditional independence for the multiple spells within each household and between households. These assumptions are also too restrictive, because multiple spells within each household tend to be much more similar than that between households.

In this paper, we propose an alternative bivariate model (a composite transition model) for analyzing the multiple-spell and multiple-destination data. Our new method not only treats both the interpurchase timing and the brand choice as random outcomes but also incorporates the dependence between successive spells for the same household. First we use the transition model approach (Diggle *et al.* 1994, pp. 192-193) in longitudinal studies to develop the likelihood as a product of conditional bivariate densities of the response in each spell given the last spell. Then we rewrite each of the conditional bivariate densities as a product of two densities by conditioning on the brand choice variable as opposed to conditioning on the purchase timing variable as in the dynamic McFadden model. We use a mixed multinomial logit model for the first density (the discrete choice transition part) and the Cox proportional hazards model (Cox 1972) for the second density (the conditional density of the interpurchase time given the two sandwiching brand choices).

Our method has several advantages over the existing approaches for analyzing multiple-spell and multiple-destination data. (1) We jointly model the bivariate outcomes of the brand choice and interpurchase time; thus the model has higher fidelity than the marginal model for either variable or the conditional model discussed earlier. (2) Our model incorporates a dependence structure across successive purchases, thus providing a more realistic framework in modeling the multiple-spell data. (3) Our model links many existing methods, for example, the marginal method for discrete choice and the conditional method by Vilcassim and Jain (1991). From the next section where detailed derivation for our likelihood is given, we see the conditional model of Vilcassim and Jain is a component of our composite model. Therefore, we are providing a theoretical justification for the conditional transition model given by Vilcassim and Jain. Consequently, our derivation helps to explain and clarify the roles that some of the existing methods play in brand-choice studies. (4) Our model has an easy interpretation where each component of the composite model (either discrete choice model or duration model) is well understood. (5) The interpurchase time part of our composite model is modeled on the minimum of the latent transition times. In this respect, it is simpler than that of Vilcassim and Jain (1991). It also avoids the pitfalls of the competing risk models (cf. Prentice *et al.* 1978). (6) In addition to the choice specific variables, we also incorporate the chooser-specific variables in the discrete choice analysis; we are not aware of any previous brand choice studies that include the chooser-specific variables in the discrete choice model. Finally, (7) Our model is easy to fit, requiring a straightforward manipulation of the data set and then applying existing software, for example, SAS or S-Plus. We can fit the discrete choice model separately from the duration model. The former requires that we divide the data into L (the number of brands) groups; each group consists of pairs of consecutive brand choices with the same past brand. Then we fit a mixed multinomial logit model to the present brand choices for each group. The latter requires that we fit separate Cox regression models for each of the subgroups of the interpurchase times with the same sandwiching transition states. This can be done easily by many software products. For example, we can use the procedure PHREG in SAS with a BY statement with two BY variables: the two consecutive states.

We apply our new method to the Nielsen household scanner panel data that was recently analyzed by Chintagunta and Prasad (1998). In the analysis, we use essentially the same predictors (price, feature, household size, previous volume, and non-detergent expenditure) that they use except we add the household size (chooser-specific) in the discrete choice model. We found that not only price and feature advertisement are significant among all the choice-specific variables, but also household size is significant among all chooser-specific variables. The effects of these predictors are generally in agreement with that in Chintagunta and Prasad. For example, from the brand choice study, reduced price and feature advertisement enhance purchasing. Moreover, we found that price is a significant factor in the timing decision in addition to household size, volume purchased earlier, and expenditure on the non-detergent products; but the effect of price on the interpurchase time is often in the opposite direction to previous findings. The magnitude of the price effect in the duration part, however, is relatively small when compared to its effect in the brand choice part. Basically, our study shows that the consumer tends to buy the detergent a little sooner as price increases. This can be explained from the market psychology that the consumers may anticipate the price will continue to rise. Given that detergent is a necessity in the household and not restricted by shelf life, it is better to buy it sooner before the price goes up again.

We organize the rest of the paper in the following way. Section 1 outlines the model formulation and discusses the discrete choice model and the duration model that are integral parts of the composite model. Section 2 discusses the estimation method. Section 3 presents the data and some numerical results, and the final section summarizes our study and remarks on future research directions.

1. MODEL FORMULATION

1.1 Transition Model

Suppose we have N households in the brand choice study. Suppose that the i^{th} household makes a total of n_i repeated purchases in a fixed period, where we record the purchase timing and the brand choice at each purchase. It is assumed that each household is allowed to choose only one of the L brands for each purchase. Let w_{ij} denote the observed calendar time of the j^{th} purchase of the i^{th} household and let y_{ij} denote the brand chosen by the household at this epoch. We set y_{ij} to be a discrete index variable with integer values ranging from 1 to L. Then we can summarize our data for the i^{th} household by a series of bivariate observations $\{(t_{ij}, y_{ij})\}$, with $j = 1, \ldots, n_i$, where $t_{ij} = w_{ij} - w_{ij-1}$, the j^{th} interpurchase time for the i^{th} household. We define $w_{i0} = 0$. Along with these bivariate outcomes, we also observe a vector of chooser-specific variables $\mathbf{x}_{ij}$ and a vector of choice-specific variables $\mathbf{z}_{ij}$, for the j^{th} purchase of the i^{th} household, where $\mathbf{z}_{ij} = (\mathbf{z}'_{ij1}, \ldots, \mathbf{z}'_{ijL})'$, with $\mathbf{z}_{ijl}$ itself a column vector of choice-specific variables for the l^{th} brand. Choice-specific variables usually change over time. Some of the chooser-specific variables take the same value across the spells for each consumer; nevertheless, we do allow them to be time-variant covariates.

Generally, the n_i bivariate outcomes, $(t_{i1}, y_{i1}), \ldots, (t_{in_i}, y_{in_i})$, from the i^{th} household exhibit correlations, because the past values $\{(t_{i1}, y_{i1}), \ldots, (t_{ij-1}, y_{ij-1})\}$ influence the present observation (t_{ij}, y_{ij}). For brevity, we define the history for the i^{th} household at the j^{th} purchase to be $H_{ij} = \{(y_{ik}, t_{ik}), k = 1, \ldots, j-1\}$. Therefore the joint density of the n_i outcomes for the i^{th} household is given by

$$f_i((t_{i1}, y_{i1}), \ldots, (t_{in_i}, y_{in_i})) = f(t_{i1}, y_{i1}) \prod_{j=2}^{n_i} f(t_{ij}, y_{ij} | H_{ij}).$$

All the densities here are also functions of the covariates $\mathbf{x}$ and $\mathbf{z}$. We suppress them in all the equations in this subsection for simpler notation.

Suppose we are willing to accept a first-order Markov property. Then the above formula (the i^{th} subject's contribution to the likelihood) can be simplified to:

$$f_i((t_{i1}, y_{i1}), \ldots, (t_{in_i}, y_{in_i})) \quad = \quad f(t_{i1}, y_{i1}) \prod_{j=2}^{n_i} f(t_{ij}, y_{ij} | t_{ij-1}, y_{ij-1}). \tag{1}$$

For the joint model $f(t_{ij}, y_{ij} | t_{ij-1}, y_{ij-1})$, we decompose it further into a composite model by conditioning on the y_{ij} variable; that is we write the joint density as a product of the marginal density of y_{ij} and a conditional density for t_{ij} given y_{ij}, both conditioning on the (t_{ij-1}, y_{ij-1}) variable. Therefore, we have

$$f(t_{ij}, y_{ij}|t_{ij-1}, y_{ij-1}) = f(y_{ij}|t_{ij-1}, y_{ij-1})f(t_{ij}|y_{ij}, t_{ij-1}, y_{ij-1}). \tag{2}$$

Then we will make the following two assumptions in modeling the composite model. First,

$$f(y_{ij}|t_{ij-1}, y_{ij-1}) = f(y_{ij}|y_{ij-1}). \tag{3}$$

This assumes that y_{ij-1} is independent of t_{ij} conditioning on y_{ij-1}. It assumes that once we know the previous brand choice, the previous interpurchase time carries no information about the current brand choice. This makes intuitive sense.

Second, we assume

$$f(t_{ij}|y_{ij}, t_{ij-1}, y_{ij-1}) = f(t_{ij}|y_{ij-1}, y_{ij}). \tag{4}$$

This assumes that the two transition states sandwiching the interpurchase time are sufficient to determine the interpurchase time. That is, the interpurchase time is independent of the previous interpurchase time given the two sandwiching states.

We note that the first-order Markov property and the above two assumptions are equivalent to a semi-Markov assumption on the sequence $\{(t_{ij}, y_{ij}), j = 1, 2, \ldots)\}$ (see, for example, Lawless and Fong 1997).

Substituting the above two assumptions (3) and (4) into the composite model (2), we derive the likelihood function for all the households from (1):

$$\prod_{i=1}^{N} f_i((t_{i1}, y_{i1}), \ldots, (t_{in_i}, y_{in_i})) = \prod_{i=1}^{N}\prod_{j=1}^{n_i} f(y_{ij}|y_{ij-1})f(t_{ij}|y_{ij-1}, y_{ij}), \tag{5}$$

where $f(y_{i1}|y_{i0}) = f(y_{i1})$ and $f(t_{i1}|y_{i0}, y_{i1}) = f(t_{i1}|y_{i1})$.

This suggests that we need to model the first-order discrete choice transitions and to model the duration given the transition states. Both can be done based on the existing literature.

We can relax the second assumption in (4) to make the interpurchase time depend on the previous interpurchase time. In this respect, we still keep the same structure of modeling the duration given transition states except we add the previous interpurchase time as a covariate in the duration model.

1.2 Discrete Choice Model

Let us first discuss the discrete choice part. Given that the previous choice only takes on one of the L values, we consider a class of models conditional on the previous choice.

Suppose $\mathbf{x}_{ij} = (1, x_{ij1}, \ldots, x_{ijk})'$ is a set of chooser specific covariates for the j^{th} purchase of the i^{th} household. Let us assume the previous choice to be l. Let $\boldsymbol{\alpha}_m^{(l)}$ denote a $k+1$ dimension column vector. We use the notation $\pi_{ij}(m|l)$ to denote $f(y_{ij} = m|y_{ij-1} = l)$. Then the generalized logit model (Agresti 1990, Hosmer and Lemeshow 1989, and McCullagh and Nelder 1989) suggests

$$\pi_{ij}(m|l) = \frac{\exp(\boldsymbol{\alpha}_m^{(l)'}\mathbf{x}_{ij})}{\sum_{m=1}^{L}\exp(\boldsymbol{\alpha}_m^{(l)'}\mathbf{x}_{ij})}.$$

We may set $\boldsymbol{\alpha}_m^{(l)'} = (0, \ldots, 0)$ for $m = l$ in the above model for identifiability purposes. The log odds ratio for choosing the choice m over n is

$$\log\left(\frac{\pi_{ij}(m|l)}{\pi_{ij}(n|l)}\right) = (\boldsymbol{\alpha}_m^{(l)'} - \boldsymbol{\alpha}_n^{(l)'})\mathbf{x}_{ij}.$$

Let $\mathbf{z}_{ijm}$ be a column vector of choice-specific variables associated with the brand choice m for the j^{th} purchase of the i^{th} household. Let $\boldsymbol{\beta}^{(l)}$ denote a column vector with the same dimension as $\mathbf{z}_{ijm}$. To model the transition discrete choice probabilities, we can consider the multinomial logit model (McFadden 1974). That is,

$$\pi_{ij}(m|l) = \frac{\exp(\boldsymbol{\beta}^{(l)'}\mathbf{z}_{ijm})}{\sum_{m=1}^{L}\exp(\boldsymbol{\beta}^{(l)'}\mathbf{z}_{ijm})}.$$

Then

$$\log\left(\frac{\pi_{ij}(m|l)}{\pi_{ij}(n|l)}\right) = \boldsymbol{\beta}^{(l)'}(\mathbf{z}_{ijm} - \mathbf{z}_{ijn}).$$

This model implies the log odds of choosing m over n depend on the distance between the values of the variables of the two brands in the comparison. If the values of a particular variable for both brands are the same, the model asserts that this variable has no influence on the choice between the alternatives m and n.

We can also formulate a mixed multinomial logit model containing both chooser-specific and choice-specific covariates (cf. Agresti 1990). For example, we define

$$\mathbf{x}_{ijm}^{*} = (\mathbf{z}_{ijm}', \delta_{m1}\mathbf{x}_{ij}', \ldots, \delta_{mL}\mathbf{x}_{ij}')', \quad m = 1, \ldots, L,$$

where δ_{mt} equals 1 when $t = m$ and 0 otherwise, and $\mathbf{z}_{ijm}$ and $\mathbf{x}_{ij}$ are defined as before. Let $\boldsymbol{\beta}^{*(l)} = (\boldsymbol{\beta}^{(l)'}, \boldsymbol{\alpha}_1^{(l)'}, \ldots, \boldsymbol{\alpha}_L^{(l)'})'$. Then $\boldsymbol{\beta}^{*(l)'}\mathbf{x}_{ijm}^{*} = \boldsymbol{\beta}^{(l)'}\mathbf{z}_{ijm} + \boldsymbol{\alpha}_m^{(l)'}\mathbf{x}_{ij}$. Thus the mixed multinomial logit model containing both choice-specific and chooser-specific variables is

$$\begin{aligned}
\pi_{ij}(m|l) &= \frac{\exp(\boldsymbol{\beta}^{(l)'}\mathbf{z}_{ijm} + \boldsymbol{\alpha}_m^{(l)'}\mathbf{x}_{ij})}{\sum_{m=1}^{L}\exp(\boldsymbol{\beta}^{(l)'}\mathbf{z}_{ijm} + \boldsymbol{\alpha}_m^{(l)'}\mathbf{x}_{ij})} \\
&= \frac{\exp(\boldsymbol{\beta}^{*(l)'}\mathbf{x}^*_{ijm})}{\sum_{m=1}^{L}\exp(\boldsymbol{\beta}^{*(l)'}\mathbf{x}^*_{ijm})},
\end{aligned} \tag{6}$$

since, for identifiability, we can take $\boldsymbol{\alpha}_m^{(l)} = (0, \ldots, 0)'$ for $m = l$. Then the log odds of choosing m over n becomes

$$\begin{aligned}
\log\left(\frac{\pi_{ij}(m|l)}{\pi_{ij}(n|l)}\right) &= \boldsymbol{\beta}^{*(l)'}(\mathbf{x}^*_{ijm} - \mathbf{x}^*_{ijn}) \\
&= \boldsymbol{\beta}^{(l)'}(\mathbf{z}_{ijm} - \mathbf{z}_{ijn}) + (\boldsymbol{\alpha}_m^{(l)'} - \boldsymbol{\alpha}_n^{(l)'})\mathbf{x}_{ij}.
\end{aligned}$$

The mixed multinomial logit model in (6) enables us to incorporate consumer characteristics, intercept, and other choice-specific variables in the McFadden model. It shares the same nice interpretations as the generalized logit model and the McFadden model discussed earlier. We note that previous applications of the McFadden model in brand-choice studies only use choice-specific variables, such as prices, indicator of feature advertisements, and indicator of special store display (see, for example, Chintagunta and Prasad 1998 and Jain *et al.* 1994); and we feel that there is no particular reason for not including chooser-specific variables, such as household size.

1.3 Duration Model

For the conditional model of the interpurchase time given the sandwiching brand choices, we apply standard models in duration analysis (Kalbfleisch and Prentice, 1980). In particular, let $h(t_{ij}|y_{ij-1} = l, y_{ij} = m)$ denote the hazard function of t_{ij} conditional on the two sandwiching brand choices. Then, we consider the proportional hazards model

$$h(t_{ij}|y_{ij-1} = l, y_{ij} = m) = h_0^{(lm)}(t)\exp(\boldsymbol{\lambda}^{(lm)'}\mathbf{x}_{ij} + \boldsymbol{\xi}^{(lm)'}\mathbf{z}_{ij}) \tag{7}$$

where $l, m = 1, \ldots, L$, $\boldsymbol{\lambda}^{(lm)}$ is a column vector of the same length as $\mathbf{x}_{ij}$ (the chooser-specific variable for the j^{th} purchase of the i^{th} household), $\boldsymbol{\xi}^{(lm)}$ is a column vector of the same length as $\mathbf{z}_{ij}$ (the choice-specific variable for the same purchase), and $h_0^{(lm)}(t)$ is the baseline hazard function for the state l to state m transition. Additionally, we consider the model with the previous interpurchase time added as part of the covariates to relax the assumption 2 in (4) we made earlier.

Therefore, the conditional density $f(t_{ij}|y_{ij-1}=l, y_{ij}=m)$ is modeled by

$$f(t_{ij}|y_{ij-1}=l, y_{ij}=m) = h(t_{ij}|y_{ij-1}=l, y_{ij}=m)\exp\left\{-\int_0^{t_{ij}} h(u|y_{ij-1}=l, y_{ij}=m)du\right\}.$$

Flinn and Heckman (1982, 1983) suggest using the Box-Cox formulation to specify the baseline hazard $h_0^{(lm)}(t)$. That is,

$$h_0(t) = \exp\left(\gamma_0 + \sum\nolimits_{k=1}^{K} \gamma_k \left(\frac{t^{\lambda_k}-1}{\lambda_k}\right)\right), \tag{8}$$

where γ_0, λ_k and γ_k, $k=1,\ldots,K$ are parameters to be estimated. These parameters depends on the interpurchase times between l to m transitions. We omit the superscripts (lm) for simplicity. One advantage of this representation of the baseline hazard is that many frequently used probability distributions for durations are nested within it. For example, if $\gamma_k = 0$ for all $k \geq 1$, then $h_0(t) = \exp(\gamma_0)$, hence the baseline hazard corresponds to an exponential distribution. Weibull and extreme value distributions are also nested within (8). Thus this formula enables researchers to test the adequacy of a particular distribution for the duration.

On the other hand, studies (for example, Jain and Vilcassim 1991) have shown that in marketing the duration time may not be able to be characterized by those common distributions nested within the Box-Cox formulation. Moreover, some common distributions are not nested in the formula. For example, lognormal and logistic distributions are not in this family. Thus the use of the Box-Cox representation is quite limited. Therefore we leave the baseline hazard totally arbitrary. The resulting model is a semiparametric Cox regression model. This model assumes that the hazard ratio (relative risk) for two interpurchase times with the same sandwiching states is a constant (with respect to t), where log of the constant is a linear function of the differences of the two covariates.

2. Estimation

Let us recall the likelihood in (5) can be written as a product of two parts

$$L = \left\{\prod_{i=1}^{N}\prod_{j=1}^{n_i} f(y_{ij}|y_{ij-1})\right\}\left\{\prod_{i=1}^{N}\prod_{j=1}^{n_i} f(t_{ij}|y_{ij-1}, y_{ij})\right\}, \tag{9}$$

where we use the mixed multinomial logit models (6) for the first part (the discrete choice transition model) and the Cox proportional hazards models (7) for the second part (the conditional densities for the interpurchase times given the two sandwiching states). To maximize this likelihood, it suffices to maximize each part separately.

Define $\mathcal{H}^{(l)} = \{(i,j)|y_{ij-1} = l\}$. It is the subset of double indices on i and j such that $y_{ij-1} = l$. Then we can rewrite the first part of the likelihood in (9) as

$$\mathcal{L}_p = \prod_{i=1}^{N}\prod_{j=1}^{n_i} f(y_{ij}|y_{ij-1}) = \prod_{l=1}^{L}\left\{\prod_{(i,j)\in\mathcal{H}^{(l)}} f(y_{ij}|y_{ij-1}=l)\right\}. \tag{10}$$

Recall our mixed multinomial logit model is defined by

$$f(y_{ij}=m|y_{ij-1}=l) = \pi_{ij}(m|l) = \frac{\exp(\boldsymbol{\beta}^{*(l)'}\mathbf{x}^*_{ijm})}{\sum_{m=1}^{L}\exp(\boldsymbol{\beta}^{*(l)'}\mathbf{x}^*_{ijm})}.$$

Therefore, the partial likelihood $\mathcal{L}_p$ can be rewritten as

$$\begin{aligned}
\mathcal{L}_p &= \prod_{l=1}^{L}\left\{\prod_{(i,j)\in\mathcal{H}^{(l)}}\prod_{m=1}^{L} f(y_{ij}=m|y_{ij-1}=l)^{I(y_{ij}=m)}\right\} \\
&= \prod_{l=1}^{L}\left\{\prod_{(i,j)\in\mathcal{H}^{(l)}}\frac{\exp(\boldsymbol{\beta}^{*(l)'}\sum_{m=1}^{L}\delta_{ijm}\mathbf{x}^*_{ijm})}{\sum_{m=1}^{L}\exp(\boldsymbol{\beta}^{*(l)'}\mathbf{x}^*_{ijm})}\right\} \\
&= \prod_{l=1}^{L} L_l(\boldsymbol{\beta}^{*(l)}),
\end{aligned} \tag{11}$$

where $\delta_{ijm} = I(y_{ij}=m)$.

Therefore, to maximize the likelihood of the discrete choice transition part, it suffices to fit each of the L mixed multinomial logit models separately, each with only transition pairs with the same brand choice of last purchase. Then we fit the mixed multinomial logit model to the present brand choice. The Newton-Raphson method can be used to maximize the likelihood $L_l(\boldsymbol{\beta}^{*(l)})$. Nevertheless, we can find the MLE easily by fitting a stratified proportional hazards model (cf. Klein and Moeschberger 1997, p. 282). For example, it can be done with the procedure PHREG with the strata option in SAS. Observe for each subject in $\mathcal{H}^{(l)}$, we can create L dummy survival times with the m^{th} subject's survival time 0 if $y_{ij} = m$ and the remaining $L-1$ subjects censored at a later time (say 1). Moreover, the covariates for the m^{th} subject are $\mathbf{x}^*_{ijm}$. Then the partial likelihood employed in the Cox regression model is exactly

$$\frac{\exp(\boldsymbol{\beta}^{*(l)'}\sum_{m=1}^{L}\delta_{ijm}\mathbf{x}^*_{ijm})}{\sum_{m=1}^{L}\exp(\boldsymbol{\beta}^{*(l)'}\mathbf{x}^*_{ijm})},$$

because the risk set for the m^{th} subjects includes all the L subjects. Then we add a strata option by ID for all subjects in $\mathcal{H}^{(l)}$. Essentially, the strata option treats each set of the L dummy survival times just created as a stratum. This will allow us to maximize $L_l(\boldsymbol{\beta}^{*(l)})$. See SAS (1995 pp. 131-144) for more details.

We apply a similar idea to fit the duration model except we need to divide our likelihood into a product of $L \times L$ terms. Recall that we use the same Cox regression model for the interpurchase times with the same sandwiching states, for example, for all the consecutive states with $y_{ij-1} = l$ and $y_{ij} = m$. Therefore, to maximize the likelihood for the interpurchase time part, it suffices to fit L^2 models separately, where each model (say the $(lm)^{th}$ model) only picks up the interpurchase times with the l to m transition. It is relatively easy to use existing software to fit separate Cox regression models for each transition to obtain the estimates of the regression coefficients and the baseline hazard function for each transition. The regression coefficients estimates are actually the maximum likelihood estimates for the partial likelihood usually employed in the Cox model that ignores the baseline hazard function. Then the baseline hazard function is estimated nonparametrically by the Breslow (1974) method.

3. DATA AND RESULTS

We reanalyze the Nielsen household scanner panel data on purchases of liquid laundry detergents in Sioux Falls, South Dakota, that was recently evaluated by Chintagunta and Prasad (1998). The observation period for these data lasted from the first week of July, 1986 to July 16, 1988, a two year span. The sample consists of 400 households who purchased only from among six national brands, Tide, EraPlus, Solo, Wisk, Surf, and All. From these households, we collected a total of 3,055 spells. For scanner data, the first and the last spells are usually incomplete. The interpurchase time of the first spell is left-truncated, and thus we also don't know the starting brand for the first spell. The last spell is usually censored at the end of the study; therefore we only know that the interpurchase time should be larger than the time observed. As a consequence, we don't know what brand the household will choose at the end of the last spell. For our data, all 400 households have incomplete first spells, and 398 households have incomplete last spells. Many of the incomplete spells have missing covariates. Special methods are needed for handling these two special situations. For our paper, however, we will simply drop these incomplete spells. The resulting complete sample consists of 320 households and 2,257 spells. Of the 320 households, most of them have less than ten spells (72%), about 50% of them have six or less spells, 90% of the households have eighteen of less spells, eight households (0.25%) have more than 30 spells, two have more than 40 spells, and one household has 87 spells.

Table 1 tabulates the frequency counts of the l to m transitions for each of the 6×6 transitions among the 2,257 spells. The table suggests the degree of competition among the brands, without controlling for the covariates. The large magnitude of the diagonal elements relative to the off diagonal ones suggests the tendency of the consumers to keep buying the same brand. In fact almost three quarters (71%) of the transitions are the repeated purchases of the same brand. Table 2 summarizes the mean interpurchase time for each transition. The mean interpurchase time for the whole sample is about 55 days.

(insert Table 1 about here)

(insert Table 2 about here)

One of the objectives of our study is to investigate the relationship between the pattern of brand switching and interpurchase times. A comparison of the empirical distributions for the interpurchase times between transitions should provide some insights. Figure 1 exhibits multiple box plots arranged in order from left to right, where each plot summarizes the interpurchase times for the $1 \rightarrow 1$, $1 \rightarrow 2$, $\ldots$, $6 \rightarrow 6$ transitions except that the $3 \rightarrow 6$ transition is missing (because of no data). An interesting finding is that the larger the frequency, the shorter the interpurchase time. This is further confirmed by Figure 2, where the average interpurchase times are plotted against the magnitude of the frequencies for each transition. A very strong curvilinear relationship is exhibited in Figure 2. This figure strongly supports the need to model the bivariate outcomes jointly. Any marginal analysis will not be able to uncover this interesting feature.

(insert Figure 1 about here)

(insert Figure 2 about here)

The choice-specific variables collected in the sample include the prices (PRICE), features advertisement (FEATURE), and special displays (DISPLAY) of the six bands. The later two are both binary variables, with 1 indicating the presence of feature advertisement or special store display at the time of purchase for each brand. The unit of the price variables is cents per ounce. Since different pack sizes of different brands exist, only the prices of the purchased size are recorded for the competing brands. The chooser-specific variables contain household size (HHSIZE), volume of detergent bought last time (STOCK), and non-detergent expenditure this trip (EXPEND). The last variable acts as proxy for "regular" and "fill-in" trips. We calculated the correlations among these variables and found that feature advertisement and special store display for some brands are highly correlated. Particularly, the values of feature advertisement are exactly the same as the values of store display for the brand ALL. Therefore, to avoid collinearity, we drop one of them in the analysis.

According to the proposed model in Section 1, we estimated 6 discrete choice models, each one with a different previous brand, and 36 Cox regression models, each one for a unique transition. The results are reported in Table 3 and Table 4 under the name of choice models and duration models, respectively. We report the results without the special display (DISPLAY) variable in both cases. In choice models, only household size from the chooser-specific variables was found to be significant. So, in the final analysis, we dropped the stock and expenditure variables. In duration models, exploratory analysis incorporating the price and feature variables of the competing brands showed them to be statistically insignificant for most of the cases. Therefore they were not included in the final estimation. Only price and feature of the brand purchased are included. Because of limited sample sizes, some of the duration models return no results or unreliable results. In these cases, we use dashes to represent the estimates. In the following, we discuss the results from Table 3 and Table 4 in more detail.

(insert Table 3 about here)

(insert Table 4 about here)

3.1 Choice Models

Examining the effect of choice-specific variables (PRICE and FEATURE) on the probability of purchasing a particular brand, we see that all estimated coefficients are of the expected signs and all of them are significant at the 0.05 level (actually all but one of them are significant at the 0.01 level). The estimated coefficients can be interpreted as in the usual logistic model. For example, reading from column 1 of Table 3 for the brand Tide, we see that an 1-cent increase in the difference between the price of any particular brand and brand Tide makes the odds of the consumer buying this brand decreases to 24% (exp(-1.424)); and a brand being on feature advertisement increases the odds of the consumer buying this brand to a factor of 5.7 (exp(1.739)).

The intercepts represent the intrinsic preference for the corresponding brands relative to the base brand. In each of the six models, we make the previous brand the base brand. From Table 3, we see that most of the estimated intercepts are negative, and among those intercepts which are significantly different from zero, all but one are negative. This reflects the feature of the transition pattern: consumers are less likely to change from one brand to another. Based on the intercept estimates, for each previous brand, we can order the brands from the most favored to the least favored. For example, we see that for consumers with a previous purchase of Tide, the first choice is TIDE, the second choice is SOLO, and then SURF, ERAPLUS, WISK, and the last ALL.

Also from Table 3, we see that the chooser-specific variable, household size, plays an non-negligible role in brand choice decision. The effects of household size on the choice of different brands are mixed. For example, for those who bought brand SURF last time, one more household member makes the odds of purchasing Tide decreases to 57%, while for those who bought SOLO last time, one more family member increases the odds of purchasing ERAPLUS to about 1.6 times.

Our findings here from the choice models agree with previous findings about brand choice (Jain *et al.* 1994) in the sense of the choice-specific variables and the consumer intrinsic preference. The inclusion of the chooser-specific variable seems to be new in brand choice studies. The real implication of the effect of household size needs more investigation.

3.2 Duration Models

Table 4 reports the estimated coefficients of chooser-specific variables and choice-specific variables from the 36 Cox regression models. The number inside the parentheses after the name of transitions are the sample size for that transition. In some of the models, we don't have enough data to estimate the coefficients or the sample sizes are too small to make the estimates reliable. In some other models, one or more variables might take the same value, thus making the coefficients of the variable or variables nonidentifiable. In these cases, we use dashes to represent the estimates.

From Table 4, we see that most of the estimated coefficients for the chooser-specific variables are of the expected signs. This is clear if we elicit those which are significant at at least 0.1 level. Household size (HHSIZE) has generally positive estimated coefficients, suggesting that an increase of household size will likely increase the rate (hazard) of making a purchase. On the other hand, volume purchased at previous occasion (STOCK) generally has negative signs which says that the more you bought last time (the more you have as stock) the less the rate of making the next purchase. The effect of non-detergent expenditure (EXPEND) on the rate of making the next purchase is not that significant in general, and for those cases where it is significant, the effects are mixed.

Turning to the choice-specific variables, we note that feature advertisement (FEATURE) is generally not significant. Price is significant for some of the models and takes positive coefficients in all of them. This implies that as the price of the brand goes up, the hazard of making the purchase increases. For example, when the transition is from WISK to SURF, a one cent increase in the price of SURF increases the odds of purchasing SURF to 1.008. The magnitude of the change in odds is not large, nevertheless, this finding of the price effect on the timing decision is an interesting departure from some previous findings (see Vilcassim and Jain 1991 and Chintagunta and Prasad 1998 among many others) where the price variable usually has the opposite effect. However, the new finding does make some intuitive sense: as the price of the brand goes up, consumers might be anticipating further price increases which encourage consumers to make the purchase sooner rather than later.

In Figure 3, we plot the estimated baseline hazard function for transition times from TIDE to TIDE. We see that the baseline hazard is nonmonotonic. In particular, it fluctuates frequently with many peak and valleys, and with piecewise local trends. This feature precludes the use of the Box-Cox formulation for the baseline hazard.

Table 5 (on p. 23) reports the results for the duration model where we include the previous interpurchase time (PREDUR) in the covariates. We see the predictor (PREDUR) is significant at the 0.05 level only in 5 out of the 22 models. The price still has the same positive sign as in Table 4 whenever significant in both tables. Table 5 reveals very similar qualitative interpretations as in Table 4. The sample size is different here because we delete the observations with missing covariates.

(insert Figure 3 about here)

(insert Table 5 about here)

4. SUMMARY AND DISCUSSION

In this paper, we developed a new method to model jointly the interpurchase time and brand choice. Our model accounts for the dependence structure across successive purchases from the same household via a first-order Markov property. For the decomposition of the joint model, we found that conditioning on the brand choice variable has more advantages than has the dynamic McFadden model where the conditioning is done on the interpurchase time. It allows us to fit the discrete choice model and duration model separately. In addition, it also allows us to fit different distributions for the duration times between different transitions and different multinomial logit models for different previous brand choices. Best of all, all these fittings can be easily done with commonly used software that includes SAS or S-Plus.

We applied our model to scanner data of detergent purchase. We found that among choice-specific variables, price and feature advertisement are significant factors for consumer's choice decisions. Household size, as a chooser-specific variable, is also important in the choice decision process. We also found that price is a significant factor in the timing decision. The effect of price on the interpurchase time is in the opposite direction to previous findings, but the magnitude of the price effect (changes in odds) to the interpurchase time is much smaller than that to the brand choice.

Data similar to the brand choice data appear also in other fields; thus our method can be applied. For example, in labor economics, data are collected on when a person changes labor status among the employed, unemployed, and out of labor force status (cf. Lancaster, 1990, and Heckman and Singer, 1985). In sociology, event history data on marriage, separation, and divorce may be collected (cf. Yamaguchi, 1991). In medicine, life history data are collected to study relapse-remitting processes (cf. Anderson *et al.*, 1993). The only difference is that repeated transition to the same previous state is not allowed. We listed three such examples in our opening discussion. In such cases, we need to modify our multinomial logit model slightly to accommodate the no-repetition transition.

Our model can be made more parsimonious by assuming some of the transitions share the same model. For example, we might assume $\boldsymbol{\beta}^{*(l)} = \boldsymbol{\beta}^{*(m)}$ if we believe the switching patterns are quite similar between the last purchase choices l or m. Similarly, we can make equal coefficients assumptions for some of the parameters in the Cox model.

Our method does not yield stable results for small sample sizes. The above discussion on collapsing models would help. Furthermore, we can employ the techniques used in small domain estimation to borrow strength from neighboring categories to treat small sample-size problems.

We have ignored the incomplete observations at the two end points of the panel data for each household. The study can be improved by adding the incomplete observations. For the beginning of the monitoring period, we can impute the state space variable at time 0. That is, we impute the brand choice for the last purchase before the monitoring period. Various imputation schemes can be explored. It would be interesting to study the effects of the imputation schemes. For the end of the study, we can include right-censored data for the time between the last purchase and the ending monitoring period. This can be easily done in each of the 36 Cox regression models.

We have adopted a first-order Markov assumption in our derivation of the likelihood. This assumption can be relaxed to incorporate higher order Markov chains (cf. Diggle *et al.*, 1994, p. 195). Although it increases the complexity of the analysis, it may be needed when the AR(1) assumption is known to be inadequate.

We have made two assumptions in the conditioning decomposition of our likelihood. First, we assume $f(y_{ij}|t_{ij}, y_{ij-1}) = f(y_{ij}|y_{ij-1})$, then we use the discrete choice model with the explanatory variables observed at the i, j epoch to model this transition state space model. With time varying covariates, it would be worthwhile to consider models with explanatory variables observed at the $i, j-1$ time. Given that this transition model should depend on the interpurchase time, we can relax our first assumption by allowing the interpurchase time t_{ij} as part of the explanatory variable. On the second assumption, we assume $f(t_{ij}|y_{ij}, t_{ij-1}, y_{ij-1}) = f(t_{ij}|y_{ij}, y_{ij-1})$. This assumption can be relaxed by adding t_{ij-1} as a covariate in the Cox regression model. Our evaluation in Table 5 reveals that the dependence on the previous interpurchase time for the scanner data is not very significant.

With respect to the duration modeling, in addition to the Cox regression model, we can also consider the log linear model, the additive hazards regression model, or the generalized additive hazards models (Hastie and Tibshirani, 1990).

Unobserved hetorogeniety among households would often affect the analysis (cf. Jain *et al.*, 1994, Chintagunta and Prasad, 1998). It would be worthwhile to incorporate a random-effects specification (frailty) into our model and compare the two results.

We have assumed the state space remained unchanged during the study. Often, a new product is introduced into the market, or a product is phased out. It would, therefore, be worthwhile to consider dynamic models and change-point models for this kind of situation.

ACKNOWLEDGMENTS

The authors wish to thank Professor Pradeep Chintagunta for providing them the source of the data set and clarifying their queries on the data set. They also wish to thank Michael P. Cohen for constructive suggestions.

References

Agresti, A. (1990), *Categorical Data Analysis,* New York: John Wiley and Sons.

Andersen, P. K., Borgan, O., Gill, R. D., and Keiding, N. (1993), *Statistical Models Based on Counting Processes,* New York: Springer-Verlag.

Breslow, N. E. (1974), "Covariance Analysis of Censored Survival Data," *Biometrics,* 30, 89-99.

Chintagunta, P. K., and Prasad, A. R. (1998), "An Empirical Investigation of the 'Dynamic McFadden' Model of Purchase Timing and Brand Choice: Implications for Market Structure," *Journal of Business & Economic Statistics,* 16, 2-12.

Cox, D. R. (1972), "Regression Models and Life-tables (with discussion)," *Journal of the Royal Statistical Society Series B,* 34, 187-220.

Diggle, P. J., Liang, K., and Zeger, S. L. (1994), *Analysis of Longitudinal Data,* Oxford, U.K.: Oxford Science Publications.

Flinn, C., and Heckman, J. (1982), "Models for the Analysis of Labor Force Dynamics." In R. Bassmann and G. Rhodes, eds., *Advances in Econometrics,* Vol. 1. Greenwich, Conn.: JAI Press.

—-(1983), "The Likelihood Function for the Multistate-Multiepisode Model in 'Models for the Analysis of Labor Force Dynamics.'" In R. Bassmann and G. Rhodes, eds., *Advances in Econometrics,* Vol. 3. Greenwich, Conn.: JAI Press.

Gönül, F., and Srinivasan, K. (1993), "Consumer Purchase Behavior in a Frequently Bought Product Category: Estimation Issues and Managerial Insights from a Hazard Function Model with Heterogeneity," *Journal of the American Statistical Association,* 88, 1219-1227.

Hastie, T. J., and Tibshirani, R.J. (1990) *Generalized Additive Models*, London: Chapman and Hall.

Heckman, J., and Singer, B. (1984), "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," *Econometrica,* 52, 271-320.

Heckman, J., and Singer, B. (1985), *Longitudinal Analysis of Labor Market Data,* Cambridge, U.K.: Cambridge University Press.

Hosmer, D. W., and Lemeshow, S. (1989), *Applied Logistic Regression,* New York: John Wiley and Sons.

Jain, D. C, and Vilcassim, N. J. (1991), "Investigating Household Purchase Timing Decisions: A Conditional Hazard Function Approach," *Marketing Science,* 10, 1-23.

Jain, D. C., Vilcassim, N. J., and Chintagunta, P. K. (1994), "A Random-Coefficients Logit Brand-Choice Model Applied to Panel Data," *Journal of Business & Economic Statistics,* 12, 317-328.

Kalbfleisch, J. D., and Prentice, R. L. (1980), *The Statistical Analysis of Failure Time Data,* New York: John Wiley and Sons.

Klein, J. P., and Moeschberger, M. L. (1997), *Survival Analysis Techniques for Censored and Truncated Data,* New York: Springer.

Lancaster, T. (1990), *The Econometric Analysis of Transition Data,* Cambridge: Cambridge University Press.

Lawless, J. F., and Fong, D. Y. T. (1997), "State Duration Models in Clinical and Observational Studies," to appear in *Statistics in Medicine.*

McCullagh, P., and Nelder, J. A. (1989) *Generalized Linear Models,* (2nd ed), London: Chapman and Hall.

McFadden, D. (1974), "Conditional Logit Analysis of Qualitative Choice Behavior," in *Frontiers of Econometrics,* ed. P. Zarembka, New York: Academic Press, pp.105-142.

Prentice, R. L., Kalbfleisch, J.D., Peterson, A.V. Jr., Flounoy, N., Farewell, V.T., and Breslow, N.E. (1978), "The Analysis of Failure Times in the Presence of Competing Risks," *Biometrics,* 34, 541-554.

SAS Institute Inc. (1995), *Logistic Regression Examples Using the SAS System, Version 6, First Edition,* Cary, NC: SAS Institute Inc.

Vilcassim, N. J., and Jain, D. C. (1991), "Modeling Purchase-Timing and Brand-Switching Behavior Incorporating Explanatory Variables and Unobserved Heterogeneity," *Journal of Marketing Research,* 28, 29-41.

Yamaguchi, K. (1991), *Event History Analysis,* Newbury Park, CA: SAGE Publications, Inc.

Table 1: Transition Matrix.

Freq (Tot Pct)			*To*				*Row*
Row Pct / Col Pct	TIDE	WISK	ERAPLUS	SURF	SOLO	ALL	*Total*
TIDE	**415** (18.4)	**59** (2.6)	**40** (1.8)	**52** (2.3)	**12** (0.5)	**10** (0.4)	588 (26.1)
	70.6/70.3	10.0/10.0	6.8/6.8	8.8/8.8	2.0/2.0	1.7/1.7	
WISK	**67** (3.0)	**461** (20.4)	**37** (1.6)	**28** (1.2)	**8** (0.4)	**3** (0.1)	604 (26.8)
	11.1/11.1	76.3/76.2	6.8/6.1	4.6/4.6	1.3/1.3	0.5/0.5	
ERAPLUS	**37** (1.6)	**32** (1.4)	**320** (14.2)	**33** (1.5)	**12** (0.5)	**0** (0.0)	434 (19.2)
	8.5/8.4	7.4/7.3	73.7/72.9	7.6/7.5	2.8/2.7	0.0/0.0	
SURF	**56** (2.5)	**43** (1.9)	**26** (1.2)	**196** (8.7)	**14** (0.6)	**4** (0.2)	339 (15.0)
	16.5/16.9	12.7/13.0	7.7/7.8	57.8/59.0	4.1/4/2	1.2/1.2	
SOLO	**11** (0.5)	**5** (0.2)	**12** (0.5)	**21** (0.9)	**171** (7.6)	**1** (0.0)	221 (9.8)
	5.0/5.0	2.3/2.3	5.4/5.5	9.5/9.6	77.4/78.1	0.5/0.5	
ALL	**4** (0.2)	**5** (0.2)	**4** (0.2)	**2** (0.1)	**2** (0.1)	**54** (2.4)	71 (3.1)
	6.6/5.6	7.0/6.9	5.6/5.6	2.8/2.8	2.8/2.8	76.1/75.0	
Col Total	590 (26.1)	605 (26.8)	439 (19.5)	332 (14.7)	219 (9.7)	72 (3.2)	2257

Table 2: Average Interpurchase Time (in Days) by Transitions.

			To				*Row*
From	TIDE	WISK	ERAPLUS	SURF	SOLO	ALL	*Average*
TIDE	49	61	67	65	117	74	54
WISK	75	52	71	70	70	146	57
ERAPLUS	54	55	53	44	46	-	53
SURF	68	81	59	45	82	55	56
SOLO	64	56	48	71	46	123	50
ALL	92	66	198	51	48	62	71
Column Average	54	55	58	52	53	67	55

Table 3: Parameter Estimates for Choice Models.

	Previous Brand					
Variables	TIDE	WISK	ERAPLUS	SURF	SOLO	ALL
INT (T)		0.195	-2.643^{a}	1.384^{a}	-2.566^{a}	-0.891
		(0.387)	(0.521)	(0.486)	(0.859)	(2.572)
INT (W)	-3.687^{a}		-3.075^{a}	0.052	-4.483^{a}	-3.176^{c}
	(0.450)		(0.522)	(0.513)	(1.254)	(1.781)
INT (E)	-3.052^{a}	-0.485		-0.067	-3.906^{a}	-3.642
	(0.539)	(0.495)		(0.609)	(0.917)	(2.808)
INT (Su)	-2.918^{a}	-1.116^{b}	-3.886^{a}		-5.787^{a}	1.299
	(0.498)	(0.536)	(0.584)		(0.957)	(1.745)
INT (So)	-2.517^{a}	-1.158	-2.625^{a}	-2.796^{a}		0.744
	(0.771)	(0.884)	(0.730)	(0.916)		(2.268)
INT (A)	-4.394^{a}	-4.000^{a}	na	-0.870	na	
	(0.866)	(1.327)	na	(1.256)	na	
HHSIZE (T)		-0.189	0.048	-0.555^{a}	-0.091	0.161
		(0.116)	(0.131)	(0.140)	(0.297)	(0.896)
HHSIZE (W)	-0.125		-0.312^{a}	-0.867^{a}	-0.179	0.466
	(0.124)		(0.145)	(0.176)	(0.439)	(0.592)
HHSIZE (E)	-0.197	-0.099		-0.213	0.457^{c}	1.525^{b}
	(0.149)	(0.151)		(0.155)	(0.253)	(0.744)
HHSIZE (Su)	-0.162	-0.481^{b}	0.076		0.829^{a}	-1.124
	(0.148)	(0.198)	(0.140)		(0.235)	(0.984)
HHSIZE (So)	-0.441	-0.432	-0.254	0.278		-0.780
	(0.269)	(0.302)	(0.216)	(0.202)		(1.116)
HHSIZE (A)	-0.860^{b}	-0.826	na	-2.011^{a}	na	
	(0.349)	(0.589)	na	(0.741)	na	
PRICE	-1.424^{a}	-1.348^{a}	-1.168^{a}	-1.418^{a}	-0.970^{a}	-1.929^{a}
	(0.117)	(0.116)	(0.127)	(0.134)	(0.187)	(0.507)
FEATURE	1.739^{a}	1.164	1.348^{a}	1.775^{a}	1.046^{b}	2.586^{a}
	(0.259)	(0.224)	(0.278)	(0.318)	(0.487)	(0.894)

T=TIDE, W=WISK, E=ERAPLUS, Su=SURF, So=SOLO, A=ALL;
The standard errors are in parentheses;

^{a} significant at 0.01 level, ^{b} significant at 0.05 level, ^{c} significant at 0.10 level.

Table 4: Parameter Estimates for Duration Models

Variable	T→T (415)	T→W (59)	T→E (40)	T→Su (52)	T→So (12)	T→A (10)
HHSIZE	0.093[a]	-0.064	0.390[b]	0.062	6.597[b]	-1.183
	(0.04)	(0.093)	(0.162)	(0.117)	(2.939)	(1.132)
STOCK	-0.007[a]	-0.003	0.003	-0.008	0.051	0.077[b]
	(0.002)	(0.005)	(0.006)	(0.006)	(0.038)	(0.034)
EXPEND	0.001	0.008	-0.009	-0.006	-0.030[b]	0.189[b]
	(0.001)	(0.005)	(0.005)	(0.006)	(0.014)	(0.089)
PRICE	0.002[a]	-0.000	0.002	-0.001	-0.012	0.867
	(0.001)	(0.002)	(0.002)	(0.002)	(0.008)	(0.909)
FEATURE	0.171	0.440	-0.088	-0.253	-8.564[b]	-
	(0.131)	(0.353)	(0.564)	(0.367)	(3.991)	-
	W→T (67)	W→W (461)	W→E (37)	W→Su (28)	W→So (8)	W→A (3)
HHSIZE	-0.120	0.258[a]	0.216	0.652[b]	3.161[c]	-
	(0.104)	(0.038)	(0.163)	(0.300)	(1.772)	-
STOCK	-0.003	-0.001	-0.007	-0.014[c]	0.053	-
	(0.004)	(0.001)	(0.008)	(0.008)	(0.052)	-
EXPEND	0.001	-0.001	-0.002	-0.006	-0.098	-
	(0.004)	(0.001)	(0.004)	(0.008)	(0.078)	-
PRICE	-0.000	0.001	-0.001	0.008[a]	0.034[c]	-
	(0.002)	(0.001)	(0.003)	(0.003)	(0.020)	-
FEATURE	-0.113	0.135	0.171	0.265	4.259	-
	(0.288)	(0.110)	(0.465)	(0.571)	(3.967)	-
	E→T(37)	E→W (32)	E→E (320)	E →Su (33)	E→So (12)	E→A (0)
HHSIZE	-0.104	-0.224	0.072[c]	0.189	0.793[c]	-
	(0.165)	(0.159)	(0.038)	(0.147)	(0.412)	-
STOCK	-0.011[b]	0.004	-0.006[a]	-0.010	0.041	-
	(0.006)	(0.006)	(0.002)	(0.007)	(0.025)	-
EXPEND	0.005	0.005	0.003	-0.003	-0.033[c]	-
	(0.005)	(0.008)	(0.002)	(0.011)	(0.019)	-
PRICE	-0.000	-0.002	0.002[a]	-0.000	0.011[b]	-
	(0.003)	(0.003)	(0.001)	(0.001)	(0.005)	-
FEATURE	0.191	0.258	-0.061	0.130	1.709	-
	(0.489)	(0.414)	(0.191)	(0.570)	(1.381)	-
	Su→T(56)	Su→W (43)	Su→E (26)	Su→Su (196)	Su→So (14)	Su→A (4)
HHSIZE	0.218[c]	0.145	0.301[b]	0.133[a]	5.811	-
	(0.119)	(0.151)	(0.140)	(0.043)	(549.8)	-
STOCK	-0.011[c]	0.002	-0.006	-0.006[b]	0.019[c]	-
	(0.006)	(0.006)	(0.012)	(0.003)	(0.011)	-
EXPEND	-0.014[c]	0.006	-0.000	0.004	0.029[a]	-
	(0.008)	(0.006)	(0.002)	(0.003)	(0.015)	-
PRICE	0.006[a]	-0.001	-0.001	0.001	-0.004	-
	(0.002)	(0.002)	(0.002)	(0.001)	(0.006)	-
FEATURE	-0.809[b]	-0.555	-0.239	-0.049	-	-
	(0.350)	(0.382)	(0.579)	(0.241)	-	-
	So→T (11)	So→W (5)	So→E (12)	So→Su (21)	So→So (171)	So→A (1)
HHSIZE	-0.243	-	0.305	0.211	-0.010	-
	(0.429)	-	(0.401)	(0.187)	(0.076)	-
STOCK	-0.008	-	0.012	0.001	-0.001	-
	(0.030)	-	(0.015)	(0.008)	(0.002)	-
EXPEND	-0.001	-	-0.029	0.007[c]	0.003[c]	-
	(0.003)	-	(0.024)	(0.004)	(0.001)	-
PRICE	0.010	-	-0.013	-0.001	0.003[b]	-
	(0.009)	-	(0.009)	(0.002)	(0.001)	-
FEATURE	1.659	-	-	-0.406	0.281	-
	(1.138)	-	-	(0.655)	(0.377)	-
	A→T (4)	A→W (5)	A→E (4)	A→Su (2)	A→So (2)	A→A (54)
HHSIZE	-	-	-	-	-	0.352[b]
	-	-	-	-	-	(0.162)
STOCK	-	-	-	-	-	-0.005
	-	-	-	-	-	(0.006)
EXPEN	-	-	-	-	-	0.001
	-	-	-	-	-	(0.005)
PRICE	-	-	-	-	-	-0.180
	-	-	-	-	-	(0.296)
FEATURE	-	-	-	-	-	-
	-	-	-	-	-	-

T=TIDE, W=WISK, E=ERAPLUS, Su=SURF, So=SOLO, A=ALL;

The standard errors are in parentheses;

[a] significant at 0.01 level, [b] significant at 0.05 level, [c] significant at 0.10 level.

Table 5: Parameter Estimates for Duration Models Controlling for Previous Duration

Variable	T→T (369)	T→W (45)	T→E (33)	T→Su (43)	T→So (7)	T→A (7)
HHSIZE	0.118[a]	-0.236[c]	0.326[c]	0.070	-	-
	(0.039)	(0.139)	(0.190)	(0.121)	-	-
STOCK	-0.007[a]	0.00	-0.005	-0.114[c]	-	-
	(0.002)	(0.006)	(0.006)	(0.006)	-	-
EXPEND	-0.000	0.012	-0.009	-0.002	-	-
	(0.001)	(0.008)	(0.010)	(0.007)	-	-
PRICE	0.002[a]	-0.002	0.003	-0.003	-	-
	(0.001)	(0.002)	(0.003)	(0.002)	-	-
FEATURE	0.077	0.095	-0.765	0.174	-	-
	(0.139)	(0.411)	(0.703)	(0.458)	-	-
PREDUR	-0.003[a]	0.002	-0.013[a]	-0.007[a]	-	-
	(0.001)	(0.003)	(0.004)	(0.002)	-	-
	W→T (54)	W→W (408)	W→E (29)	W→Su (20)	W→So (7)	W→A (1)
HHSIZE	-0.095	0.223[a]	-0.051	0.991[a]	-	-
	(0.128)	(0.040)	(0.226)	(0.383)	-	-
STOCK	-0.005	-0.000	-0.005	-0.020[b]	-	-
	(0.005)	(0.002)	(0.010)	(0.008)	-	-
EXPEND	-0.001	-0.001	0.002	-0.169[c]	-	-
	(0.004)	(0.001)	(0.008)	(0.009)	-	-
PRICE	-0.000	0.001	-0.002	0.011[b]	-	-
	(0.002)	(0.001)	(0.004)	(0.005)	-	-
FEATURE	-0.220	0.096	0.242	1.162	-	-
	(0.346)	(0.118)	(0.539)	(0.892)	-	-
PREDUR	-0.001	-0.002[c]	-0.002	-0.034	-	-
	(0.004)	(0.001)	(0.004)	(0.019)	-	-
	E→T(35)	E→W (25)	E→E (284)	E →Su (23)	E→So (10)	E→A (0)
HHSIZE	-0.154	-0.045	0.030	0.017	0.460	-
	(0.165)	(0.237)	(0.041)	(0.221)	(0.408)	-
STOCK	-0.015[b]	-0.001	-0.006[a]	0.004	0.024	-
	(0.006)	(0.008)	(0.002)	(0.009)	(0.026)	-
EXPEND	0.008	0.003	0.003	-0.007	-0.043[b]	-
	(0.006)	(0.008)	(0.002)	(0.012)	(0.022)	-
PRICE	-0.001	-0.000	0.002[b]	0.001	0.003	-
	(0.003)	(0.003)	(0.001)	(0.001)	(0.007)	-
FEATURE	0.287	0.332	0.163	-1.094	-0.249	-
	(0.497)	(0.594)	(0.206)	(0.814)	(1.685)	-
PREDUR	-0.004	0.010	-0.008[a]	-0.007	-0.006[b]	-
	(0.004)	(0.007)	(0.002)	(0.005)	(0.005)	-
	Su→T(47)	Su→W (34)	Su→E (21)	Su→Su (168)	Su→So (13)	Su→A (2)
HHSIZE	0.017	-0.014	0.522[b]	0.145[a]	5.429	-
	(0.150)	(0.187)	(0.265)	(0.048)	(633.9)	-
STOCK	-0.013[b]	0.002	-0.009	-0.006[b]	0.013	-
	(0.006)	(0.006)	(0.014)	(0.003)	(0.012)	-
EXPEND	-0.005	0.007	0.000	0.003	0.028[c]	-
	(0.009)	(0.006)	(0.002)	(0.003)	(0.015)	-
PRICE	0.005[c]	0.000	-0.003	0.001	-0.006	-
	(0.002)	(0.002)	(0.003)	(0.001)	(0.007)	-
FEATURE	-0.697	-0.493	-0.377	-0.227	-	-
	(0.466)	(0.448)	(0.697)	(0.279)	-	-
PREDUR	-0.004	-0.003	0.005	-0.003	-0.008	-
	(0.005)	(0.004)	(0.008)	(0.002)	0.009	-
	So→T (10)	So→W (5)	So→E (11)	So→Su (17)	So→So (149)	So→A (1)
HHSIZE	-	-	0.173	0.163	0.026	-
	-	-	(0.427)	(0.232)	(0.086)	-
STOCK	-	-	-0.002	0.005	0.002	-
	-	-	(0.025)	(0.011)	(0.002)	-
EXPEND	-	-	-0.024	0.012	0.002	-
	-	-	(0.025)	(0.008)	(0.002)	-
PRICE	-	-	-0.012	-0.001	0.003[b]	-
	-	-	(0.009)	(0.002)	(0.001)	-
FEATURE	-	-	-	-0.260	0.380	-
	-	-	-	(0.004)	(0.386)	-
PREDUR	-	-	-0.028	0.004	-0.010	-
	-	-	(0.035)	(0.007)	(0.003)	-
	A→T (3)	A→W (3)	A→E (2)	A→Su (1)	A→So (2)	A→A (48)
HHSIZE	-	-	-	-	-	0.273
	-	-	-	-	-	(0.192)
STOCK	-	-	-	-	-	-0.005
	-	-	-	-	-	(0.006)
EXPEN	-	-	-	-	-	0.000
	-	-	-	-	-	(0.005)
PRICE	-	-	-	-	-	-0.138
	-	-	-	-	-	(0.307)
FEATURE	-	-	-	-	-	-
	-	-	-	-	-	-
PREDUR	-	-	-	-	-	-0.004
	-	-	-	-	-	(0.003)

Figure 1: Side by Side Box Plots of Interpurchase Times for the 6×6 Transitions.

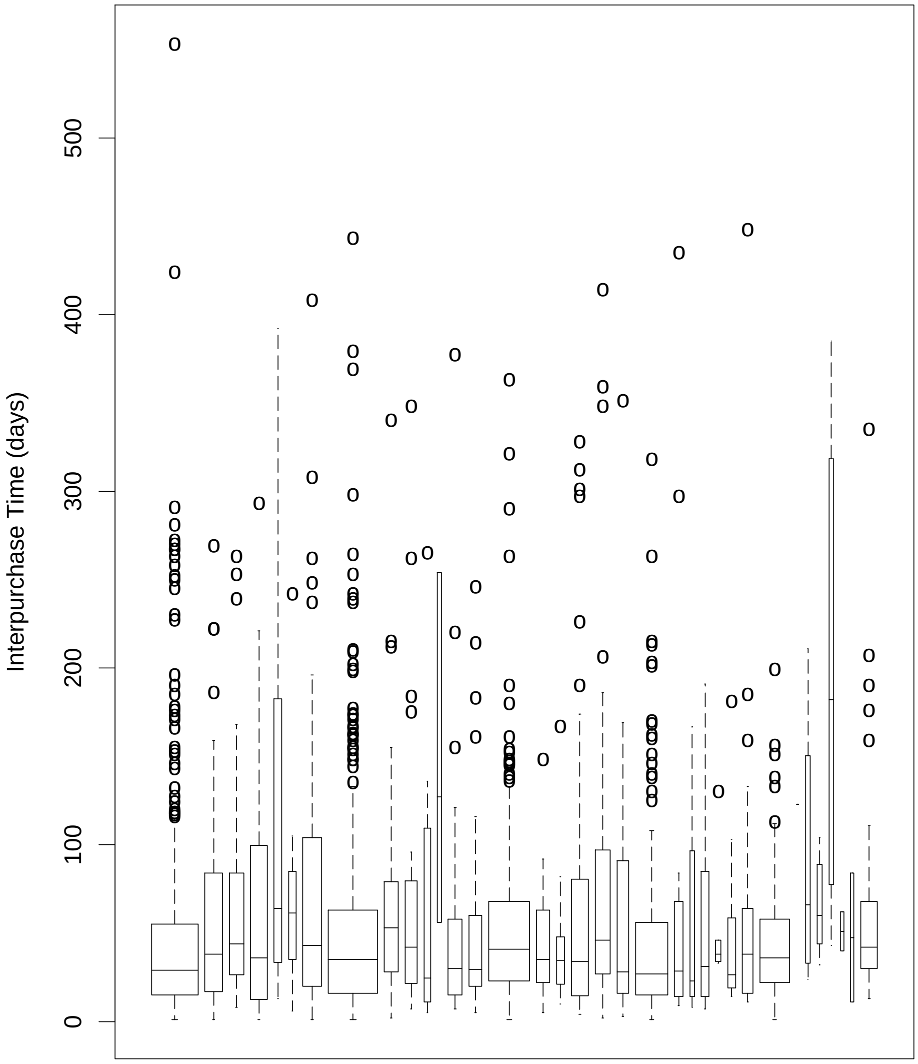

Figure 2: Average Interpurchase Time vs Transition Frequencies.

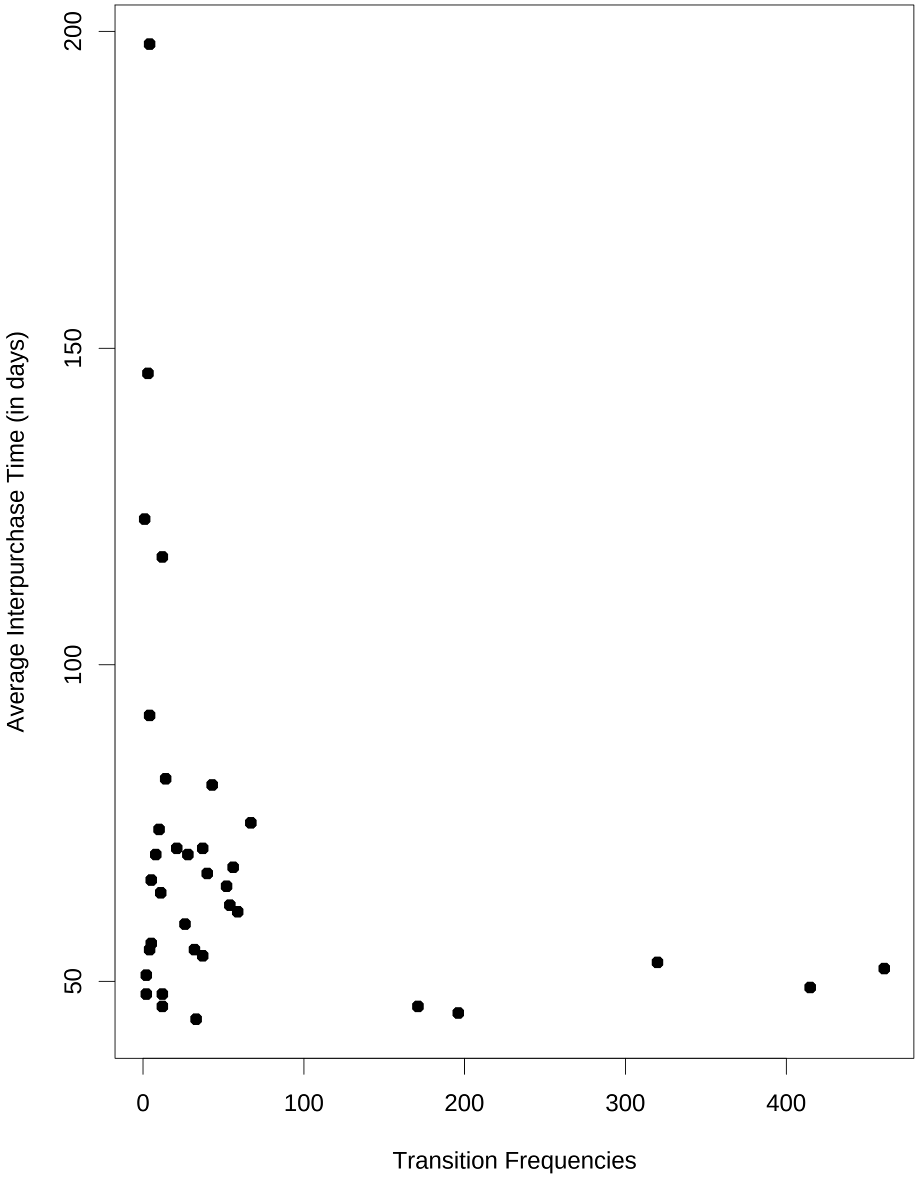

Figure 3: Estimated Baseline Hazard Function for Transition TIDE→TIDE.