

A Continuous-Time Model of Departure Time Choice for Urban Shopping Trips

Chandra R. Bhat and Jennifer L. Steed

Department of Civil Engineering, The University of Texas at Austin

ECJ 6.806, Austin, Texas 78712, Tel: (512) 471-4535, Fax: (512) 475-8744

E-mail: bhat@mail.utexas.edu, jlsteed@mail.utexas.edu

ABSTRACT:

This paper proposes a continuous-time hazard duration model for urban shopping trip departure time choice. The time frame for the analysis of departure time is the entire day. The continuous-time model uses a non-parametric baseline hazard distribution, employs a non-parametric representation for the time-varying effect of covariates, and accommodates time-varying covariates. These econometric issues are important to recognize and consider in a departure time model spanning the entire day. The model also accommodates unobserved heterogeneity and recognizes the "rounding" of reported departure times by individuals in surveys to an integral multiple of five minutes. The continuous-time model is estimated using shopping trip data from the 1996 activity survey collected in the Dallas-Fort Worth metropolitan area. The effects of individual and household socio-demographics, employment attributes, and trip-related characteristics on departure time choice are presented and discussed. Application of the continuous-time model to forecast temporal shifts in shopping trip-making due to changes in socio-demographic characteristics and trip-chaining behavior is demonstrated. The use of the formulation as a powerful evaluation tool to manage and influence traffic patterns through dynamic congestion pricing control schemes is also highlighted.

1. INTRODUCTION

The time-of-day choice of individuals for trip-making is an important determinant of the temporal pattern of traffic on urban roadways. An analysis of individual time-of-day choice is, therefore, important for a) planning the development and construction of new transportation infrastructure to serve forecasted temporal demands, b) examining the potential responses to improved operational strategies (such as advanced traffic control strategies and real-time information provision), and c) assessing the effectiveness of time-of-day specific transportation demand management measures (such as congestion pricing).

The importance of studying departure time choice has been recognized for some time now, and several studies have examined this choice behavior in the past couple of decades. Much of this earlier literature has been directed toward examining commuter departure time decisions as a function of individual/household demographic characteristics, work-related attributes, transportation system characteristics, and real-time information provision regarding traffic delays (see, for example, Abkowitz, 1981; Mannering, 1989; Chin, 1990; Hendrickson and Plank, 1984; Mahmassani and Jou, 1996; Small, 1982; and Mannering and Hamed, 1990). The emphasis in the past on commuter departure time choice is a consequence of the commute periods representing the most congested time of the day. More recently, some researchers (see, for example, Kumar and Levinson, 1995; Hunt and Patterson, 1996; Bhat, 1998a, 1998b and Steed and Bhat, 2000) have examined departure time decisions in the context of non-work trips. These researchers point to the near vacuum of studies on non-work trip departure time and identify at least two reasons for analysis in this area. First, non-work trips constitute an increasingly large proportion of urban trips in general and the peak periods in particular. Thus, traffic congestion cannot be simply attributed to commute trips. Second,

individuals are likely to have more temporal flexibility for pursuing non-work activities than they do for commuting to work. The implication of this greater flexibility for non-work trips is that socio-demographic changes and/or transportation control measures may have a more significant impact on such trips than on the less flexible work trips.

The objective of this paper is to contribute toward improving the techniques of analysis of departure time decisions in the context of non-work trips. In almost all earlier studies of non-work departure time choice, time has been treated as a discrete variable; that is, departure time alternatives are represented by several temporally contiguous discrete time periods such as early morning, a.m. peak, a.m. off-peak, p.m. off-peak, p.m. peak, and evening. The choice among these alternatives is modeled using a discrete choice model. There are four inter-related limitations of such a discrete choice representation of departure time. First, the use of discrete time periods requires the rather *ad hoc* partitioning of the day into time intervals; unfortunately, different temporal partitioning schemes can, and in general will, lead to different model results. Second, the discrete choice structure treats time points near the boundaries of intervals as belonging to one or the other of the aggregate time periods. In reality, two closely-spaced time points on either side of a discrete interval boundary are likely to be perceived as being similar rather than as distinct alternatives. Third, a discrete structure provides departure time of trips only in aggregate intervals. The consequent loss in time resolution precludes the determination of temporal traffic patterns at the fine level that may be needed for the design of traffic control strategies or for air quality modeling (see Steed and Bhat, 2000 for a detailed discussion). Of course, one may address the need for fine temporal resolution within the context of a discrete structure by identifying a sufficiently large number of very small discrete time intervals collectively spanning the entire day. This, however, would imply the estimation of a very high

number of parameters and, more fundamentally, would defeat the purpose of using a discrete choice structure. Fourth, the use of a discrete structure requires the retention of the same aggregate intervals for forecasting as used in estimation. This would be an impediment to examining potential future control and pricing strategies during time periods which are narrower than (or simply different from) those used in estimation.

The current paper uses a continuous-time approach to modeling departure time. Such an approach overcomes the limitations of the discrete structure indicated above. A continuous-time representation has been used in the past to model departure time decisions, but studies using such a representation have been confined to narrow time periods of the day. For example, Hamed and Mannering (1993) and Bhat (1998b) examine non-work trip departure time for workers after they arrive home from work at the evening; Abu-Eisheh and Mannering (1989) examine morning commute departure time; and Mannering and Hamed (1990) study the length of time commuters delay their departure from work in the late afternoon to avoid peak period congestion. The structure for continuous-time representations within such narrow time periods of the day can take the form of a proportional hazard duration model, which assumes that the effect of covariates on departure time (or, more precisely, on the hazard of termination of the duration preceding trip departure) does not change over time. However, assuming fixed effects of covariates is untenable when the time domain for consideration is the entire day, as is the case in the current research. For example, the effect of children on the termination of the activity duration preceding participation in recreational activity may be much more "accelerated" during the evening than in earlier times of the day because the evening is most convenient (from schedule considerations) for joint activity participation with children. Such sudden non-monotonic accelerations (or decelerations) in the effect of variables over

the course of the day cannot be captured by the traditional proportional hazard or accelerated lifetime models (the accelerated lifetime model allows time-varying effects, but specifies the time-varying effects to be monotonic and smooth in the time domain). Similarly, previous studies of continuous departure time choice focusing on narrow time periods have not had to deal with changing covariates over time. But time-varying covariates must be accommodated in a departure time choice model spanning the entire day.

To summarize, this paper formulates and applies a continuous-time hazard duration model for trip timing decisions which accommodates time-varying coefficients as well as time-varying covariates. In addition, the model recognizes the presence of unobserved characteristics affecting departure time decisions. The model uses a non-parametric baseline hazard distribution which does not impose any *a priori* structure on the duration distribution. Finally, the proposed model recognizes the “rounding-off” in reported departure times to the nearest five minute multiple of clock time.

The continuous-time methodology is applied to analyze the departure time for shopping trips of adult individuals (greater than 16 years of age) using data from an activity survey conducted in the Dallas-Fort Worth area. Since a vast majority of individuals who made a shopping trip during the survey day made only one such trip, we further focus in this paper only on departure time modeling for the first shopping trip of the day.

The remainder of this paper is structured as follows. The next section presents the model structure. Section 3 provides an overview of the data source and variables used in the empirical analysis. Section 4 discusses the empirical results. Section 5 demonstrates the application of the model. The final section summarizes important findings.

2. MODEL STRUCTURE

Let T be a non-negative random variable representing the duration (in minutes) from midnight until the departure time of an individual for a shopping trip (we suppress the index for the individual in this presentation). The hazard at time u on the continuous time scale, $\lambda(u)$, is defined as the instantaneous probability that the duration preceding shopping trip departure will end in an infinitesimally small time period h after time u , given that the duration has not elapsed until time u . A precise mathematical definition for the hazard in terms of probabilities is as follows:

$$\lambda(u) = \lim_{h \rightarrow 0^+} \frac{\Pr(u \leq T < u + h \mid T > u)}{h}. \quad (1)$$

This mathematical definition immediately makes it possible to relate the hazard to the density function $f(\cdot)$ and cumulative distribution function $F(\cdot)$ for T . Specifically, since the probability of the duration terminating in an infinitesimally small time period h after time u is simply $f(u) \times h$, and the probability that the duration does not elapse before time u is $1 - F(u)$, the hazard rate can be written as:

$$\lambda(u) = \frac{f(u)}{[1 - F(u)]} = \frac{f(u)}{S(u)} = \frac{dF/du}{S(u)} = \frac{-dS/du}{S(u)} = \frac{-d \ln S(u)}{du}, \quad (2)$$

where $S(u)$ is a convenient notational device which will be referred as the endurance (or survivor) probability and which represents the probability that the duration did not end prior to u (*i.e.*, that the duration "endured" until time u)¹.

¹The concept of a hazard rate originated in the Industrial Engineering and Biometrics fields and has been used in those fields for several decades now (see, for example, Berkson and Gage, 1952, and Goodman, 1953). Its use in the economics and transportation fields has been relatively recent, though much of the new

To accommodate the effect of exogenous covariates and the effect of unobserved attributes, we use the functional form shown below:

$$\lambda(u) = \lambda_0(u) \exp(\beta_u' x_u) w, \quad (3)$$

where $\lambda_0(u)$ is the baseline hazard, x_u is a vector of covariates specific to time u (some or all of whose elements may be time-varying), β_u is a corresponding vector of coefficients, and w is an unobserved heterogeneity term that is assumed to be independent of x_u . The form in equation (3) is more general than the proportional hazard (PH) form and collapses to the PH form if $\beta_u = \beta \forall u$ (McCall, 1994).

The shape of the baseline hazard function, $\lambda_0(u)$, in equation (3) has important implications for duration dynamics. One may adopt a parametric shape or a non-parametric shape for the baseline hazard (see Bhat, 2000 for an extensive discussion). A problem with the parametric approach is that it will, in general, inconsistently estimate the hazard function when the assumed parametric form is incorrect (Meyer, 1990). The advantage of using a non-parametric form is that, even when a particular parametric form is appropriate, the resulting estimates are consistent and the loss of efficiency (resulting from disregarding information about the hazard's distribution) may not be substantial (Meyer, 1987). In the current paper, we use a non-parametric form for the hazard. In the non-parametric approach, the duration scale is split into several smaller grouped intervals. Assuming a constant hazard (*i.e.*, an exponential duration distribution) within each grouped interval, one can estimate the continuous-time step-function hazard shape. The reader will note that this grouping of

advances in hazard models has emerged in these fields (see Kiefer, 1988, Hensher and Mannering, 1994 and Bhat, 2000 for reviews of hazard-based models in the context of the economics and transportation fields).

the time scale is not inconsistent with an underlying continuous process for the duration data. In fact, the grouping may be motivated from considerations of “rounding off” in the reporting of underlying continuous duration times and the need for accounting for the resulting tied nature of departure time data (Bhat, 1996).

The form of the time-varying coefficient vector β_u as a function of time (in equation 3) may also be parametric or non-parametric. In most cases, it is difficult to identify an appropriate continuous parametric form for the influence of covariates over time. Therefore, we adopt a non-parametric form for the effect of covariates. In this method, the overall time duration is sliced into several grouped intervals within which the effect of covariates is assumed to be constant, but across which the pattern of variation can be of any nature to be determined from empirical data.

To proceed further, we define k as the index for grouped intervals along the time scale ($k=1,2,3,\dots,K$). Define a^k as the upper bound of the grouped interval k . Then, using equations (2) and (3), the probability that an individual's duration preceding departure to shop terminates in interval k (*i.e.*, the probability that the individual departs to shop in time period k) may be written conditional on the unobserved heterogeneity term w as:

$$\begin{aligned} \text{Prob}[t=k] | w &= \text{Prob}[a^{k-1} < T \leq a^k] | w = S(a^{k-1}) | w - S(a^k) | w \\ &= \exp\left[-\int_0^{a^{k-1}} w \lambda_0(u) \exp(\beta'_u x_u) du\right] - \exp\left[-\int_0^{a^k} w \lambda_0(u) \exp(\beta'_u x_u) du\right]. \end{aligned} \quad (4)$$

Assuming that β_u and x_u are constant within time periods k (*i.e.*, $\beta_u = \gamma_k$ and $x_u = z_k$ for all $u \in \{a^{k-1}, a^k\}$), and using the notation $\lambda_0(u) = \exp(\eta_k)$ for all $u \in \{a^{k-1}, a^k\}$ based on the assumption of a constant hazard within intervals, the expression above may be written as:

$$Prob[t = k] | w = \exp\left[-\sum_{j=0}^{k-1} w \Delta_j \exp(\eta_j + \gamma'_j z_k)\right] - \exp\left[-\sum_{j=0}^k w \Delta_j \exp(\eta_j + \gamma'_j z_j)\right], \quad (5)$$

where Δ_k is the length of time interval k , and $\eta_0 = -\infty$ and $\eta_K = +\infty$ by notational convention. The expression in (5) collapses to a proportional hazard form if and only if $\gamma_k = \gamma \forall k$.

To determine the unconditional probability that the individual departs for the shopping trip in time period k , we integrate over the distribution $G(\cdot)$ of w :

$$Prob[t = k] = \int_0^{\infty} \exp\left[-\sum_{j=0}^{k-1} w \Delta_j \exp(\eta_j + \gamma'_j z_j)\right] dG(w) - \int_0^{\infty} \exp\left[-\sum_{j=0}^k w \Delta_j \exp(\eta_j + \gamma'_j z_j)\right] dG(w), \quad (6)$$

There are two possible approaches to specifying the $G(\cdot)$ distribution. One approach is to assume a parametric mixing form and the second is to use a non-parametric form in which the distribution is replaced by a finite number of support points with associated probability masses. In the first approach, a convenient and commonly used parametric distribution is the gamma distribution since it provides a closed-form expression for the unconditional probability. This is the distribution we use in the current study. We also attempted a non-parametric approach but encountered convergence problems in the implementation of the approach.

If w is distributed as a gamma random variable with a mean one (a normalization) and variance σ^2 , the unconditional probability of failure in the discrete time period k can be written using the moment-generating function properties of the gamma distribution as (see Johnson and Kotz, 1970):

$$Prob[t = k] = \left[1 + \sigma^2 \left\{ \sum_{j=0}^{k-1} \Delta_j \exp(\eta_j + \gamma'_j z_j) \right\}\right]^{-\sigma^{-2}} - \left[1 + \sigma^2 \left\{ \sum_{j=0}^k \Delta_j \exp(\eta_j + \gamma'_j z_j) \right\}\right]^{-\sigma^{-2}}. \quad (7)$$

The parameters to be estimated in the model include the $\boldsymbol{\eta}_k$ and $\boldsymbol{\gamma}_k$ vectors ($k = 1, 2, \dots, K$) and the variance σ^2 of the gamma mixing distribution. The likelihood function for estimation of these parameters may be written after introducing the index q for individuals as:

$$\mathcal{L} = \prod_{q=1}^Q \prod_{k=1}^K \left\{ \left[1 + \sigma^2 \left\{ \sum_{j=0}^{k-1} \Delta_j \exp(\boldsymbol{\eta}_j + \boldsymbol{\gamma}'_j z_j) \right\} \right]^{-\sigma^{-2}} - \left[1 + \sigma^2 \left\{ \sum_{j=0}^k \Delta_j \exp(\boldsymbol{\eta}_j + \boldsymbol{\gamma}'_j z_j) \right\} \right]^{-\sigma^{-2}} \right\}^{M_{qk}},$$

where

$$M_{qk} = \begin{cases} 1 & \text{if individual } q's \text{ trip begins in period } k \\ 0 & \text{otherwise} \end{cases} \quad (q = 1, 2, \dots, Q, k = 1, 2, \dots, K). \quad (8)$$

The continuous-time non-parametric baseline hazard is obtained from the $\boldsymbol{\eta}_k$ parameters as

$$\lambda_0(u) = \exp(\boldsymbol{\eta}_k) \text{ for all } u \in \{a^{k-1}, a^k\}.$$

Maximization of the log-likelihood function is achieved using the GAUSS matrix programming language. The analytic gradients of the log-likelihood function with respect to the parameters have been coded.

3. THE DATA

3.1 Data Source and Sample Description

The data source used for this analysis is the 1996 activity survey conducted in the Dallas-Fort Worth metropolitan area by the North Central Texas Council of Governments (NCTCOG). This survey included an activity diary to be filled out by all members of the household. The activity diary collected information on all activities undertaken during the diary day. For non-travel activities, information on the activity type, start and end times of participation, and location was collected. For

travel activities, information on the mode of travel used, costs incurred, and trip duration was collected. In addition, the survey elicited individual and household socio-demographic information.

The shopping trip sample used in the current analysis was obtained from the activity survey file after a series of data assembly and consistency screening procedures. The final sample comprises 1315 individuals with a shopping trip during the day. Table 1 provides descriptive information on shopping departure time. The column titled "Risk Set" provides information on the number of individuals who are "at risk" of termination of the "before-shopping" duration in period k ; that is, it provides the number of individuals who are yet to depart for their shopping trip at the beginning of period k . The column labeled "# Depart" indicates the number of individuals who depart for their shopping trip in period k . The continuous-time sample hazard is easily computed from equation (5) by ignoring the covariate effects and the unobserved heterogeneity term and using the sample share that departs in period k in place of the probability of departure in period k :

$$\lambda_k = -\frac{1}{\Delta_k} \ln \left[1 - \frac{\text{Failures}_k}{\text{Risk Set}_k} \right]. \quad (9)$$

The sample hazard values and their t-statistics are shown in the last two columns of Table 1. The continuous-time sample hazard is also plotted in Figure 1. The hazard function is non-monotonic, though there is a general increasing trend over time. That is, the likelihood of departing to shop tends to increase over time among those who have not yet pursued their shopping activity.

3.2 Variable Specification

The choice of variables for potential inclusion in the model was guided by previous departure time choice research and by intuitive arguments regarding the effect of exogenous variables on shopping departure time choice. Four broad sets of variables were considered: individual socio-demographics, household socio-demographics, individual employment-related attributes, and trip-related characteristics.

Individual socio-demographic characteristics explored in our specifications included dummy variables for sex, ethnicity and education level, and continuous representations of income and age. Household socio-demographic characteristics considered in the model included household size, the number and age distribution of children, the number and employment status of household adults, and the continuous value of household income. Individual employment-related attributes included dummy variables for employed individuals, self-employed individuals, students, homemakers and retired persons, as well as a continuous variable indicating the number of hours worked per week. The final category of variables was the trip-related characteristics, including whether the trip is a home-based or a non-home based trip, whether the trip is undertaken for grocery or non-grocery shopping, and the trip travel time and trip travel cost.

The effect of individual and household demographics, and employment-related attributes, on the hazard was allowed to vary over time. Specifically, the effect of these variables was allowed to vary across six discrete intervals in the day: early morning (midnight-6:30 a.m.), a.m. peak (6:30 a.m.-9:00 a.m.), a.m. off-peak (9:00 a.m.-noon), p.m. off-peak (noon-4:00 p.m.), p.m. peak (4:00 p.m.-6:30 p.m.), and evening (6:30 p.m.-midnight).

The time-varying covariates, such as trip travel times and costs, are available only for peak and off-peak periods of the day for the Dallas-Fort Worth region. These times and costs are applied to the six intervals identified above based on whether the time period is off-peak or peak.

We arrived at the final specification based on a systematic process of eliminating variables found to be insignificant in previous specifications and based on considerations of parsimony in representation.

4. EMPIRICAL RESULTS

The empirical results are discussed in three sections. The next section presents the baseline hazard estimates. Section 4.2 discusses the covariate effects. Finally, section 4.3 interprets the results associated with unobserved heterogeneity.

4.1 Baseline Hazard

The baseline hazard function for the estimated model is shown in Figure 2. This baseline hazard has the same overall shape as the sample hazard in Figure 1. There is a general positive duration dependence in the hazard function; that is, the longer an individual waits to participate in shopping, the more likely s/he is to participate in shopping in the next time period. The baseline hazard, like the sample hazard, exhibits non-monotonic behavior.

While the general shape is the same between the sample and the baseline hazards, there are also important differences between the two. Specifically, compared to the sample hazard, the baseline hazard is smaller until about 9 a.m. in the morning, larger between 9 a.m. and 4 p.m., and again smaller thereafter. The change in the hazard function is also less abrupt from one period to the

next in the baseline hazard compared to the sample hazard, and there is a more consistent positive duration dependence in the baseline hazard. This is to be expected, since the sharp "rises" and "drops" in the sample hazard can be attributed, in part, to the covariate effects.

4.2 Covariate Effects

In this section, we discuss the effects of covariates on the departure time hazard. It should be noted that a positive coefficient on a covariate implies that the covariate increases the departure time hazard rate (equation 3) or, equivalently, the covariate "hastens" departure to shop. The magnitude of the effect of the covariates may be assessed by computing the percentage change in the hazard due to a one unit change in the covariate. From the functional form in (3), this can be written as: $\{\exp(\beta_{ul}) - 1\} \times 100$ for a one unit change in x_{ul} , the l th covariate at time u .

Table 2 shows the estimated covariate effects for the final model specification. As indicated earlier, our formulation permits time-varying effects of covariates. One can empirically test the restricted proportional-hazard model that constrains the effect of covariates to be the same over time with the more general formulation used here. The log-likelihood value of our formulation is -4345, while that of the proportional hazard formulation is -4430. A comparison of the two formulations using a likelihood ratio test provides a test value of 170, which is larger than the chi-squared table statistic with 4 degrees of freedom at any reasonable level of significance. Thus, the proportional hazard formulation is rejected in favor of our formulation. In general, for most econometric modeling, it would appear that the proportional hazard form would be very restrictive. At the least, it is appealing to estimate a model of time-varying covariate effects and test the proportional hazard assumption. The next few sections discuss the covariate effects by variable group.

4.2.1 Individual socio-demographic variables

Among the individual socio-demographics, the effect of age indicates that older individuals are likely to pursue shopping later in the day than younger individuals. Non-linear effects of age on the departure time hazard were also tested, but did not improve data fit substantially and were also difficult to interpret.

The sign of the coefficient on female sex indicates a higher departure hazard for women during the p.m. off-peak and p.m. peak periods; that is, women are more likely than men to participate in shopping during these periods. Similarly, Caucasians appear to be less likely to participate in shopping during the p.m. peak. These effects may represent the higher tendency of female workers and non-Caucasian workers to pursue shopping on the return home from work than male workers and Caucasians, respectively.

The final individual socio-demographic variable is income. The results imply that individuals with high income have a lower hazard (lower departure rate) during the off-peak and evening time periods, possibly reflecting tighter work schedule constraints. Non-linear specifications of the effect of income on the departure time hazard were tested, but did not improve the model significantly.

4.2.2 Household socio-demographics

The only household socio-demographic variables appearing in the final model specification are those associated with the presence and age distribution of children. The results show that individuals whose households have young children (less than or equal to 5 years of age) are less likely to participate in shopping during the mid-day (a.m. and p.m. off-peak) hours. This may be a result of the need to take care of the biological needs of young children at home during the mid-day.

The presence of children between the ages of 6 and 11 increases the hazard of the before-shopping duration during the am peak times; that is, presence of children in this age group results in a greater likelihood of participation during the a.m. peak period.

4.2.3 Individual Employment-Related Attributes

The individual employment-related attributes in the final model specification include externally (non-self) employed status, a dummy variable indicating if the individual is self-employed, student status, and another dummy variable indicating if the individual is a retiree.

The effect of these variables is quite intuitive. Externally employed individuals are less likely to depart for shopping during the mid-day period relative to unemployed individuals due to work-schedule constraints. A similar result holds for self-employed individuals and students compared to unemployed individuals. A comparison of the coefficients for the mid-day period on the externally employed, self-employed, and student status variables indicates that self-employed individuals and students are more likely to participate in shopping activities during the mid-day than externally employed individuals. Perhaps, this is a reflection of lesser schedule rigidity for self-employed individuals and students, because of which they are more easily able to “sandwich” a shopping activity during lunch time than externally employed individuals. Also, self-employed individuals appear to be more likely than other individuals to pursue shopping during the late evening period.

The effect of retirement status suggests that retirees stay away from the morning peak period and participate in shopping primarily during the p.m. peak period.

4.2.4 Trip-Related Attributes

The type of shopping trip has a substantial impact on departure time. As one would expect, trips to shop during the morning and a.m. peak periods tend to be home-based trips, while those in the p.m. periods tend to be non-home based trips pursued from work and/or as part of a series of errands in a single sojourn from home. The results also indicate that grocery shopping trips are less likely to be pursued in the middle of the day than non-grocery trips.

We also estimated models with separate baseline hazard functions for a) home-based grocery trips, b) non-home based grocery trips, c) home-based non-grocery trips, and d) non-home based non-grocery trips. However, the simpler specification adopted in the current paper was found to be adequate (from a data fit standpoint) in capturing differences in the hazard among the different shopping trip types.

A very interesting finding from the analysis is that travel time and travel cost do not appear to impact departure time choice for shopping trips. We attempted several interaction effects of level-of-service variables with type of trip and with individual characteristics such as sex and income. But none of these effects were statistically significant. Thus, level-of-service variables do not appear to play a role in shopping trip departure time choice. Steed and Bhat (2000) found a similar result in an earlier discrete choice formulation for departure time. This lack of effect of level-of-service on departure time choice may be due to two related reasons. First, shopping activities are organized around work constraints and other household schedule considerations, and are pursued at the most convenient time within these schedule considerations. Level-of-service variables may therefore not influence shopping departure decisions. Second, shopping trips tend to be of a rather short length.

In the context of the shorter lengths of shopping trips, there is likely to be little variation in level-of-service across time periods. This may be manifesting itself in the form of the lack of any level-of-service effect on shopping departure time decisions.

4.3 Unobserved Heterogeneity

Unobserved heterogeneity is included in the model using a gamma mixing distribution. The unobserved heterogeneity term acts multiplicatively on the hazard function. The variance of this term is estimated to be 0.3125; the t-statistic is 2.21. Thus, the results indicate the presence of factors other than those included in the model that impact departure times. Such factors may include inherent individual preferences for certain departure times, household responsibilities/interactions not captured by the household variables included in the model, and other miscellaneous unobserved determinants. It is well-established now that failure to control for unobserved heterogeneity can produce severe bias in the nature of duration dependence and the estimates of the covariate effects (Heckman and Singer, 1984; Lancaster, 1985).

The magnitude of the variance of the unobserved heterogeneity term provides important information regarding the fraction of the variation in the departure hazard rate explained by the covariates and by unobserved factors (see Jain and Vilcassim, 1991). This is most easily observed by taking the logarithm of both sides of equation (3), yielding the following equation:

$$\ln \lambda(u) = \ln \lambda_0(u) + \beta'_u x_u + \ln(w). \quad (10)$$

Since the baseline hazard, $\lambda_0(u)$, is the same across observations, the variance across observations of the (log) departure time hazard is simply the variance across observations of the $\beta'_u x_u$ term (that

is, the variance due to observed heterogeneity) and the variance across observations due to unobserved factors (that is, the variance due to unobserved heterogeneity):

$$\text{Var}[\ln\lambda(u)] = \text{Var}[\beta_u'x_u] + \text{Var}[\ln(w)]. \quad (11)$$

The fraction of variation in the departure time hazard explained by unobserved heterogeneity may then be computed as $\text{Var}[\ln(w)]/(\text{Var}[\beta_u'x_u] + \text{Var}[\ln(w)])$. The variance term, $\text{Var}[\ln(w)]$, can be estimated since the variance of w is known and w is gamma-distributed. To accomplish this, one can use an approximate analytic formula (see Johnson and Kotz, 1970) or use a simulation method. We use the latter method here.

Table 3 shows the percentage variation in departure time hazard explained by unobserved heterogeneity for each of the six broad time periods in which the covariates and their effects are allowed to vary. The results show that the highest percentage variation captured by unobserved heterogeneity is in the morning period; this is not surprising since very few covariates are introduced in this period due to the sparse nature of data. The percentage variation captured by unobserved heterogeneity is higher for the p.m. peak and evening periods compared to the other non-morning periods. This might be an indication of the greater flexibility in departure time choices during the later periods of the day due to fewer schedule constraints. Overall, the results show that the hazard model performs quite well, with observed factors explaining about 65-87% of the overall variation in the log departure time hazard.

5. MODEL APPLICATION

The model formulated in this paper can be applied to obtain the temporal changes in trip-making due to changes in trip-chaining patterns, employment patterns, and household and individual socio-demographics. This is particularly relevant today because of changing trip-making patterns and shifting demographic and employment-related characteristics of the population. For instance, trip-chaining is becoming more and more prevalent today (see Misra and Bhat, 2000). Similarly, the number of employed individuals in the household and the number of self-employed individuals are on the rise (U.S. Census Bureau, 1999a). The structure of the household is also changing rapidly with an increase in households with no children (see U.S. Census Bureau, 1999b). At the same time, the population is aging. All of these changes will have an effect on departure time, and the departure time model in this paper can be used to assess these impacts and provide reliable information regarding the temporal distribution of trips for input to transportation policy and air quality analysis.

Unlike discrete choice models of departure time choice, the current hazard-based duration model is able to provide the temporal re-distribution of trips due to changes in determining variables at any desired level of temporal precision. In particular, once estimated, the model can be used in forecasting mode using equation (7) with any desired definition of time intervals. There is no need for the time intervals to correspond to the grouped intervals used in estimation; the grouping of intervals in estimation was undertaken to estimate a non-parametric baseline hazard and to accommodate the tied nature of departure time.

In the current analysis, level-of-service variables did not turn out to be significant determinants of departure time. However, this result may not be the same in a different empirical

context. For example, the lack of sensitivity to level-of-service may be a result of little variation in times and costs across time periods in the sample from the Dallas-Fort Worth region. In other metropolitan areas, where there are substantial time/cost differences across time periods (for example, because of congestion-pricing controls which are absent in the Dallas-Fort Worth area), the level-of-service variables may have a significant impact. In such a case, the proposed hazard-based formulation can provide the temporal shifts in trip-making due to varying pricing strategies across different times of the day. In the case of discrete choice models, the pricing strategies that can be evaluated are restricted to those implemented in one of the discrete periods used in estimation (for example, a.m. peak or p.m. peak). Within the broad time periods, the price cannot vary and the pricing strategy should be one that changes abruptly at the boundary points of the discrete time periods (for example, from zero to the peak period congestion price). However, a more realistic pricing strategy may be one that begins before the peak periods, gradually increases during the peak period, and falls toward the end of the peak period and after. The continuous-time hazard formulation developed in the current paper allows the analysis of congestion pricing at any time-of-day and at any pricing level rather than restrictive pricing schemes implemented only during pre-defined times of the day. It can be used as a powerful evaluation tool to manage and influence traffic patterns through dynamic pricing control strategies implemented with Intelligent Transportation Systems Technology. A limitation of the continuous-time model formulated here, however, is that it is not able to accommodate departure time changes to earlier times of the day due to implementation of a control strategy at a particular time of the day.

In this paper, we demonstrate the application of the model by focusing on the effect of two changes. The first is an increase in the “chaining” of shopping stops with other stops (leading to an

increase in non-home based shopping trips), and the second is an increase in the proportion of self-employed individuals.

The increase in trip-chaining of shopping stops is "implemented" by randomly selecting 25% of current home-based shopping trips in the sample and "converting" them into non-home-based trips. The original distribution of home-based and non-home based trips in the sample is as follows: 661 (50.3%) home-based trips and 654 (49.7%) non-home-based trips. After "implementing" the increase in trip chaining, the new distribution is 497 (37.8%) home-based trips and 818 (62.2%) non-home-based trips.

The increase in self-employed individuals is reflected by reducing the number of unemployed individuals who are non-students and non-retired individuals by 25%, and designating them as self-employed individuals. The original number of unemployed individuals and self-employed individuals are 571 (43.4%) and 193 (14.7%), respectively. With the change, the number of unemployed and self-employed individuals becomes 428 (32.5%) and 236 (17.9%), respectively.

The temporal pattern of trips before and after the changes is obtained by sample enumeration. The probability of departing in each 15-minute time interval for each individual-trip is predicted using equation (7), and the probabilities are added across trips to obtain the estimated number of departures for each of the 15-minute time periods. However, for presentation simplicity, we aggregate the results and provide the temporal patterns in six broad time periods: early morning (midnight-6:30 a.m.), a.m. peak (6:30 a.m.-9:00 a.m.), a.m. off-peak (9:00 a.m.-noon), p.m. off-peak (noon-4:00 p.m.), p.m. peak (4:00 p.m.-6:30 p.m.), and evening (6:30 p.m.-midnight).

Table 4 shows the results. The effect of an increase in trip-chaining is provided in the second column of the table. The results indicate fewer trips being made in the early morning and a.m. peak periods, and more trips during the p.m. off-peak and p.m. peak periods. This is consistent with the results in Table 2 which shows that non-home based trips tend to be pursued during the p.m. off-peak and p.m. peak periods. The decrease in the evening period is a natural consequence of the high departure rate during the p.m. periods.

The effect of an increase in self-employment is shown in the third column of the table. There is no change in the number of trips during the morning and a.m. peak periods because the hazard rates do not change during these periods (see Table 2). However, there is a decrease in trips during the off-peak periods since self-employed individuals have work schedule constraints which reduce their likelihood of shopping during the mid-day compared to unemployed individuals. Finally, the increase in trip-making during the p.m. peak and evening periods may be a reflection of the inclination of self-employed individuals (most of whom work at home) to leave home for a brief period after a day of work.

6. CONCLUSIONS

This paper proposes the use of a continuous-time hazard duration model to analyze departure time choice for shopping trips. The time frame for analysis is the entire day. The distinguishing features of the formulation are that it uses a non-parametric baseline hazard, employs a non-parametric representation for the time-varying effect of covariates, and accommodates time-varying covariates. In addition, unobserved heterogeneity is introduced using a parametric gamma distribution. Another appealing characteristic of the formulation is that it recognizes the “rounding”

of reported departure time by individuals to an integral multiple of five minutes (*e.g.*, 5,10,15,30,60 minutes, *etc.*), and accommodates the resulting “ties” in reported departure times.

The empirical analysis in the paper uses the 1996 activity survey data collected by the North Central Texas Council of Governments (NCTCOG) in the Dallas-Fort Worth area. Several sets of variables are considered in the model specifications, including individual and household socio-demographics, employment-related attributes and trip-related characteristics. A statistical test of time-constant effects of the variables on the departure time hazard rate with the time-varying formulation used in the current paper clearly indicates the superior fit of the latter over the former. This indicates a rejection of the commonly-used proportional hazard formulation.

The empirical results indicate the strong influence of individual and household socio-demographics, individual employment-related attributes, and trip-chaining behavior on shopping trip departure time choice. This has substantial implications for transportation policy analysis and for transportation-air quality analysis, especially at a time when socio-demographics and employment characteristics are changing rapidly in the population, and trip-chaining is becoming more and more prevalent. The continuous-time model developed in this paper can be used to evaluate the impacts of these changes on temporal patterns of trip-making at any desired level of temporal resolution.

An interesting finding from our model is that level-of-service characteristics do not appear to substantially impact departure time for shopping trips. This confirms the earlier finding by the authors using a discrete choice formulation. The result has significant implications for time-of-day specific transportation control measures (such as peak-period pricing or converting a general-purpose lane to a high occupancy vehicle use lane). Specifically, our results imply that there will be little to no temporal displacements of shopping trips due to TCMs. This suggests that shopping trips may

not be as temporally flexible as one might think. That is, an individual considers participation in shopping activities within the larger spectrum of daily activities that need to be pursued based on her/his individual circumstances and household structure characteristics. Thus, scheduling issues may be so overpowering in shopping participation that they render individuals insensitive to level-of-service changes. However, it should be noted that these results may be specific to the Dallas-Fort Worth area, and may be the result of the limited variation in time and costs across different times of the day in the data used for analysis. Further exploration of the effect of level-of-service variables with richer time/cost information by time of day and using data from other metropolitan areas would be valuable.

The model can be applied in a straight-forward manner in forecasting mode, as demonstrated by the application to evaluate an increase in trip-chaining of shopping stops with other stops and an increase in the percentage of self-employed individuals. Integration of such a continuous-time forecasting system of departure time choice with a traffic micro-simulation platform would be of considerable value in understanding the effects of shifts in socio-demographics and transportation system attributes on temporal traffic patterns on a network.

ACKNOWLEDGMENTS

This research was funded by the U.S. Department of Transportation through the Southwest Region University Transportation Center. The first author would also like to acknowledge funding support from the National Science Foundation grants DMS 9208758 and DMS 9313013 to the National Institute of Statistical Sciences (NISS). We are grateful to Mahmoud Ahmadi, Ken Cervenka, and Gustavo Baez of the North Central Texas Council of Governments for providing the

data and assisting with data-related issues. Two anonymous reviewers provided useful comments and suggestions on an earlier version of the paper. Lisa Weyant helped with typesetting and formatting.

REFERENCES

- Abu-Eisheh, S.A., and F.L. Mannering (1989). Discrete/Continuous Analysis of Commuters' Route and Departure Time Choices, *Transportation Research Record 1138*, 27-34.
- Abkowitz, M.D. (1981). An Analysis of the Commuter Departure Time Decision, *Transportation*, 10, 283-297.
- Berkson, J. and R.P. Gage (1952). Survival Curve for Cancer Patients Following Treatment, *Journal of the American Statistical Association*, 47, 501-515.
- Bhat, C.R. (1996). A Hazard-Based Duration Model of Shopping Activity with Nonparametric Baseline Specification and Nonparametric Control for Unobserved Heterogeneity, *Transportation Research B*, 30, 189-207.
- Bhat, C.R. (1998a). Accommodating Flexible Substitution Patterns in Multi-Dimensional Choice Modeling: Formulation and Application to Travel Mode and Departure Time Choice, *Transportation Research B*, 32, 7, 455-466.
- Bhat, C.R. (1998b). A Model of Post-Home Arrival Activity Participation Behavior, *Transportation Research B*, 32, 6, 387-400.
- Bhat, C.R. (2000). Duration Modeling, forthcoming in *Handbook of Transport*, Kenneth Button and David Hensher, eds.,
- Chin, A.T.H. (1990). Influences on Commuter Trip Departure Time Decisions in Singapore, *Transportation Research A*, 24, 5, 321-333.
- Goodman, L.A., (1953). Methods of Measuring Useful Life of Equipment Under Operational Conditions, *Journal of the American Statistical Association*, 48, 503-530.

- Hamed, M.M. and F.L. Mannering (1993). Modeling Travelers' Postwork Activity Involvement: Toward a New Methodology, *Transportation Science*, 27, 4, 381-394.
- Heckman, J. and B. Singer (1984). A Method for Minimizing the Distributional Assumptions in Econometric Models for Duration Data, *Econometrica*, 52, 271-320.
- Hendrickson, C., and E. Plank (1984). The Flexibility of Departure Times for Work Trips, *Transportation Research A*, 18, 1, 25-36.
- Hensher, D.A., and F.L. Mannering (1994). Hazard-Based Duration Models and Their Application to Transport Analysis, *Transport Reviews*, 14, 1, 63-82.
- Hunt, J. D., and D. M. Patterson (1996). A Stated Preference Examination of Time of Travel Choice for a Recreational Trip, *Journal of Advanced Transportation*, 30, 3, 17-44.
- Jain, D.C. and N.J. Vilcassim (1991). Investigating Household Purchase Timing Decisions: A Conditional Hazard Function Approach, *Marketing Science*, 10, 1, 1-23.
- Johnson, N. and S. Kotz (1970). Distributions in Statistics: Continuous Univariate Distributions, John Wiley, New York.
- Kiefer, N.M. (1988). Economic Duration Data and Hazard Functions, *Journal of Economic Literature*, 27, June, 646-679.
- Kumar, A., and D. Levinson (1995). Temporal Variations on Allocation of Time, *Transportation Research Record 1493*, TRB, National Research Council, Washington, D.C., 118-127.
- Lancaster, T. (1985). Generalized Residuals and Heterogenous Duration Models With Applications to the Weibull Model, *Journal of Econometrics*, 28, 1, 155-169.

- Mahmassani, H.S. and Jou, R.C. (1996). Bounded Rationality in Commuter Decision Dynamics: Incorporating Trip Chaining in Departure Time and Route Switching Decisions, presented at the Conference on Theoretical Foundations of Travel Choice Modeling, Stockholm, Sweden.
- Mannering, F.L. (1989). Position Analysis of Commuter Flexibility in Changing Routes and Departure Times, *Transportation Research B*, 23, 1, 53-60.
- Mannering, F.L. and M.M. Hamed (1990). Occurrence, Frequency, and Duration of Commuters' Work-to-Home Departure Delay, *Transportation Research B*, 24, 2, 99-109.
- McCall, B.P. (1994). Testing the Proportional Hazards Assumption in the Presence of Unmeasured Heterogeneity, *Journal of Applied Econometrics*, 9, 321-334.
- Meyer, B.D. (1987). Semiparametric Estimation of Duration Models, Ph.D. Thesis, MIT, Cambridge, Massachusetts.
- Meyer, B.D. (1990). Unemployment Insurance and Unemployment Spells, *Econometrica*, 58, 4, 757-782.
- Misra, R. and C.R. Bhat (2000). A Descriptive Analysis of Non-Worker Daily Activity Travel Patterns, submitted to *Transportation Research Record*, July 1999.
- Small, K.A. (1982). Scheduling of Consumer Activities: Work Trips. *The American Economic Review*, 72, 3, 467-479.
- Steed, J. and C.R. Bhat (2000). On Modeling the Departure Time Choice for Home-Based Social/Recreational and Shopping Trips, submitted to *Transportation Research Record*.

U.S. Census Bureau (1999a). Working at Home. Population Division, Journey to Work and Migration Statistics Branch,

<http://www.census.gov/population/www/socdemo/workathome/wkhtab1.html>.

U.S. Census Bureau (1999b). Projected Number of Families with Children Under 18 Years by

Type: 1995 to 2010. Population Division, Population Projections Branch,

<http://www.census.gov/population/www/projections/nathh.html>.

LIST OF TABLES

Table 1. Shopping Departure Time Distribution and Continuous-Time Sample Hazard

Table 2. Estimated Covariate Effects

Table 3. Percentage Variation in Departure Time Hazard Explained by Unobserved Heterogeneity

Table 4. Predicted Impacts of an Increase in Trip-Chaining and in Self-Employed Individuals

**Table 1. Shopping Departure Time Distribution
and Continuous-Time Sample Hazard**

Period	Clock Time	Risk Set	# Depart	Hazard	t-stat
1	12:00 AM-6:30 AM	1315	11	0.00002	3.32
2	6:30 AM-7:00 AM	1304	9	0.00023	3.00
3	7:00 AM-7:30 AM	1295	17	0.00044	4.12
4	7:30 AM-8:00 AM	1278	15	0.00039	3.87
5	8:00 AM-8:30 AM	1263	16	0.00042	4.00
6	8:30 AM-9:00 AM	1247	23	0.00062	4.80
7	9:00 AM-9:30 AM	1224	45	0.00125	6.71
8	9:30 AM-10:00 AM	1179	45	0.00130	6.71
9	10:00 AM-10:30 AM	1134	51	0.00153	7.14
10	10:30 AM-11:00 AM	1083	47	0.00148	6.86
11	11:00 AM-11:30 AM	1036	53	0.00175	7.28
12	11:30 AM-12:00 PM	983	39	0.00135	6.24
13	12:00 PM-12:30 PM	944	52	0.00189	7.21
14	12:30 PM-1:00 PM	892	68	0.00264	8.24
15	1:00 PM-1:30 PM	824	39	0.00162	6.24
16	1:30 PM-2:00 PM	785	66	0.00293	8.12
17	2:00 PM-2:30 PM	719	43	0.00206	6.56
18	2:30 PM-3:00 PM	676	41	0.00209	6.40
19	3:00 PM-3:30 PM	635	50	0.00273	7.07
20	3:30 PM-4:00 PM	585	52	0.00310	7.21
21	4:00 PM-4:30 PM	533	68	0.00455	8.24
22	4:30 PM-5:00 PM	465	64	0.00494	7.99
23	5:00 PM-5:15 PM	401	26	0.00447	5.10
24	5:15 PM-5:30 PM	375	46	0.00872	6.78
25	5:30 PM-5:45 PM	329	15	0.00311	3.87
26	5:45 PM-6:00 PM	314	48	0.01106	6.92
27	6:00 PM-6:15 PM	266	30	0.00798	5.47
28	6:15 PM-6:30 PM	236	41	0.01272	6.39
29	6:30 PM-6:45 PM	195	24	0.00876	4.90
30	6:45 PM-7:00 PM	171	37	0.01625	6.07
31	7:00 PM-7:15 PM	134	19	0.01019	4.35
32	7:15 PM-7:45 PM	115	33	0.01127	5.72
33	7:45 PM-8:15 PM	82	23	0.01097	4.77
34	8:15 PM-8:45 PM	59	28	0.02145	5.20
35	8:45 PM-9:15 PM	31	15	0.02205	3.80

Table 2. Estimated Covariate Effects

Variable	Parameter Estimate	t-statistic
Individual Socio-demographics		
Age (divided by 100)		
p.m. off-peak, p.m. peak, evening	1.429	4.95
Female		
p.m. off-peak, p.m. peak	0.264	3.00
Caucasian		
p.m. peak	-0.345	-2.27
Individual income (in thousands)		
a.m. off-peak, p.m. off-peak	-0.025	-1.60
evening	-0.090	-2.20
Household Socio-demographics		
Presence of children <= 5 years		
a.m. off-peak, p.m. off-peak	-0.333	-2.25
Presence of children 6-11 years		
a.m. peak	0.750	3.01
Employment-related Attributes		
Externally (non-self) employed		
a.m. off-peak, p.m. off-peak	-1.206	-11.23
Self-employed		
a.m. off-peak, p.m. off-peak	-0.524	-3.24
evening	0.792	2.23
Student		
a.m. off-peak, p.m. off-peak	-0.550	-2.98
Retired		
a.m. peak	-0.756	-1.92
p.m. peak	0.713	3.10
Trip-related Attributes		
Home-based trip		
morning and a.m. peak	0.848	3.59
p.m. off-peak and p.m. peak	-0.476	-5.73
Grocery shopping		
a.m. off-peak, p.m. off-peak	-0.279	-3.22

Table 3. Percentage Variation in Departure Time Hazard Explained by Unobserved Heterogeneity

Period	% variation explained by unobserved heterogeneity
Morning	35.2
a.m. Peak	22.6
a.m. Off-Peak	14.2
p.m. Off-Peak	13.3
p.m. Peak	28.2
Evening	27.4

**Table 4. Predicted Impacts of an Increase in Trip-Chaining
and in Self-Employed Individuals**

Time Period	Percentage change in number of trips during each time period due to an increase in...	
	Trip-Chaining	Self-Employed Individuals
Morning	-28.0	0.0
a.m. Peak	-18.0	0.0
a.m. Off-Peak	0.4	-9.6
p.m. Off-Peak	7.9	-4.5
p.m. Peak	4.1	6.5
Evening	-15.5	11.8

LIST OF FIGURES

Figure 1. Continuous-Time Sample Hazard Function

Figure 2. Baseline Hazard Function

Figure 1. Continuous-Time Sample Hazard Function

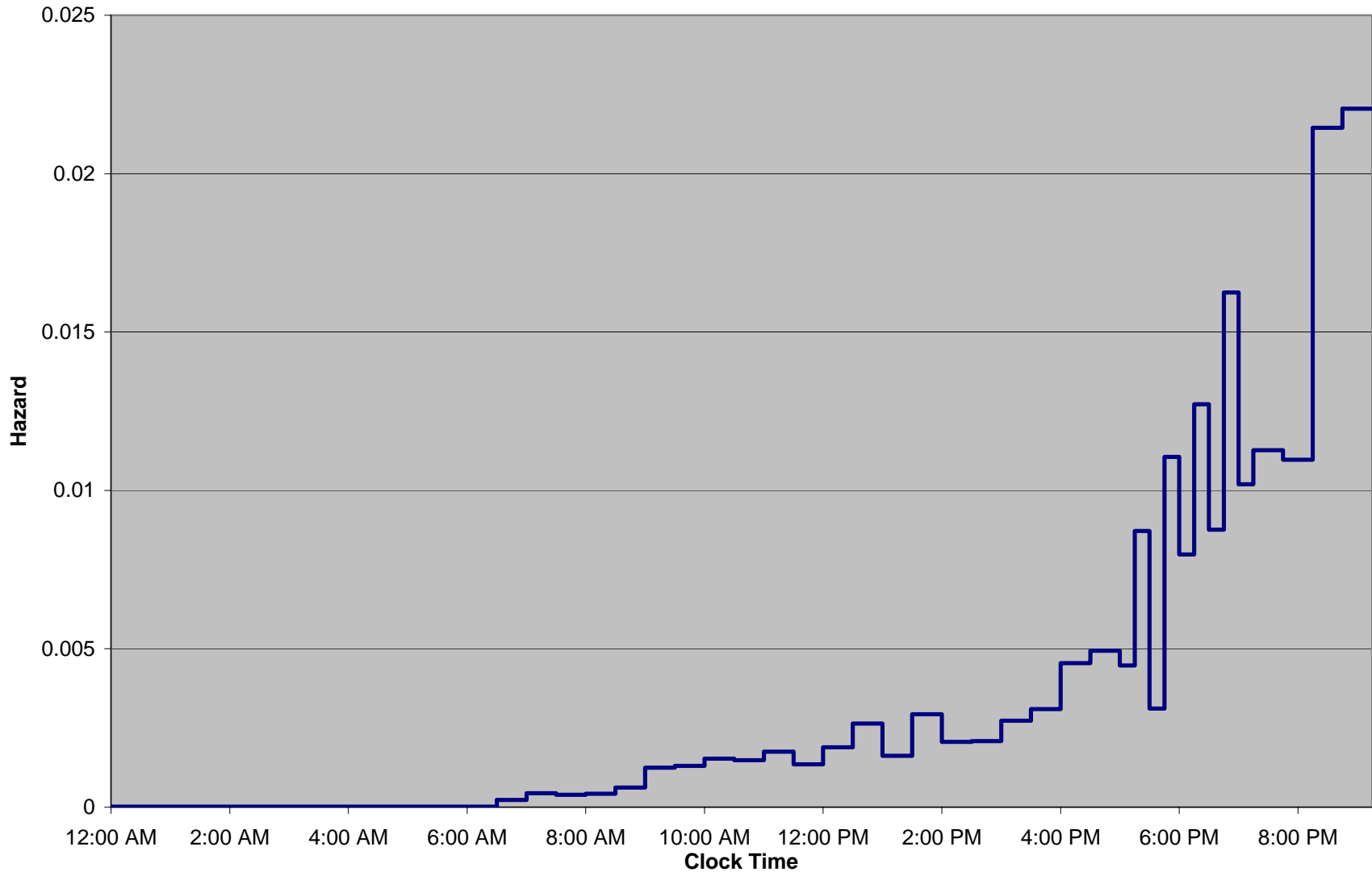


Figure 2. Baseline Hazard Function

