Continuous Departure Time Models: A Bayesian Approach

Shashank Gadda Credit
Policy/Risk Analyst HSBC
Card Services Salinas,
California
Email: shashankgc@gmail.com

Kara M. Kockelman
(Corresponding Author)
William J. Murray Jr. Associate Professor of Civil,
Architectural and Environmental Engineering The
University of Texas at Austin
ECJ 6.9, Austin, Texas 78712,
Tel: (512) 471-4379
FAX: (512) 475-8744
Email: kkockelm@mail.utexas.edu

Paul Damien
The B.M. Rankin Jr. Professor of Business
McCombs School of Business
The University of Texas at Austin
Email: Paul.Damien@mccombs.utexas.edu

Abstract:

Models of departure time presently rely on discrete choice or simple proportions across blocks of time within a 24-hour day. Duration models allow for more flexible specifications to explain both unimodal (one-peak) and multi-modal (multi-peak) data, which are common in (aggregate) departure time data. This paper offers Bayesian estimates of continuous departure time models using accelerated failure time (AFT) specifications for various trip purposes with several distributional specifications, including the lognormal, Weibull, Weibull (with and without unobserved heterogeneity), and a mixture of normals. The home-based work (HBW) and non-home -based (NHB) trip models are modeled using unimodal distributions, while the home-based non-work (HBNW) trip departure times are modeled via a bi-modal distribution. The results indicate that a Weibull with unobserved heterogeneity performs well among unimodal distributions, and that multi-peak profile can be modeled well with a mixture of normals.

Keywords: Departure time, Bayesian estimation, duration models, Weibull distribution

1. Introduction

An individual faces several choices in making a trip, namely mode, destination, departure time and route. These choices have been researched in depth and are treated explicitly in travel models. The time of day aspect generally is dealt with by applying simple proportions across blocks of time or by calibrating logit models for several times of day (typically treated as independent alternatives). Calibration is performed using maximum likelihood estimation or other, simpler techniques. Duration models, in which time is treated as a continuous variable, offer several advantages over discrete alternatives. They predict trip departure times on a continuous scale, thus enhancing the inputs to, and outputs of, models of traffic and emissions. Such model predictions influence policies targeting congestion management and air quality. In addition, continuous-time models allow one to avoid problems of temporal aggregation and period association, providing more fluid estimates of choice and illuminating finer adjustments in traveler behavior, while offering the necessary inputs for dynamic traffic assignment models. This paper looks into estimation of continuous departure time models using Bayesian techniques. This paper applies such models to the case of Austin, Texas' 1996 Travel Survey data, to predict departure times for travelers engaged in various trip types. The following presents a detailed literature review of previous departure time models. Sections 3, 4 and 5 describe the model specifications, their estimation methods, and method of comparison. Section 6 discusses the empirical findings, and Section 7 offers several conclusions.

2. Literature review

Most of the early departure time studies used multinomial logit (MNL) models for work trips. For example, Abkowitz (1981), Small (1982) and McCafferty and Hall (1982) modeled departure and arrival time choices of individuals using demographic variables (like income and age) and chosen mode. Their results suggested important effects of work schedule flexibility, income, age, occupation, and transportation-system level of service on departure time choice. Hendrickson and Plank (1984) modeled mode and departure time choice simultaneously, using an MNL model with 28 alternatives. These represented combinations of 4 modes (drive alone, auto, shared ride, and transit) and 7 different departure time intervals of 10 minutes each. Their results suggested that travelers were more likely to change their departure time than their mode. Chin (1990) also modeled morning departure times for Singapore commuters using MNL and nested logit (NL) models. The NL model was used to moderate certain violations of the independence of irrelevant alternatives (IIA) property (McFadden (1978), Ben-Akiva & Lerman (1985)). He found that departure time choice was influenced by journey time (with longer journeys requiring earlier start times, as anticipated), and that occupation and income affected one's propensity for switching departure times.

All above-mentioned studies have examined departure time for work trips. In contrast, Bhat (1998) and Steed and Bhat (2000) modeled departure time for home-based recreational and shopping trips. Bhat (1998) adopted a nested structure with mode choice at the higher level of the hierarchy and departure time choice at the lower level. A MNL form was used for mode choice; and an ordered generalized extreme value (OGEV) form, which recognizes the natural (temporal) ordering of the departure time alternatives, was adopted for departure time choice. "The OGEV structure generalizes the MNL structure by allowing an increased degree of sensitivity (due to excluded exogenous factors) between adjacent departure time alternatives compared to between non-adjacent departure time alternatives." (Bhat, 1998. pg. 362).

The MNL-OGEV model was applied to data obtained from the 1990 San Francisco Bay Area travel survey and found to perform better than MNL and NL models. Moreover, the MNL and NL models lead to biased estimates of travel time and cost coefficients and lead to inappropriate policy evaluations of congestion pricing and other transportation control measures. Steed and Bhat (2000) estimated discrete departure time choice models for four categories of HBNW trips using the 1996 activity survey data collected in the Dallas-Fort Worth region. They focused on the effects that individual and household demographics, employment attributes, and trip characteristics have on departure time choice. MNL and OGEV model structures (as in Bhat 1998) were used for modeling departure time choice among six discrete periods. In all four trip-purpose cases, the MNL model was adequate for capturing controlled effects, when compared to the OGEV results. Trip level-of-service variables had relatively low effects, suggesting "that departure times for non-work trips are not as flexible as one might expect, and are confined to certain times of day because of overall scheduling constraints." (Steed and Bhat, 2000. pg 152)

More recently, Abou-Zeid et al. (2006) and Popuri et al. (2008) sought to control for variations in network conditions (speeds and travel times) by time of day. They discretized time into half-hour intervals, for activity-based modeling applications across a variety of trip types. While their standard logit specifications neglect the ordering of time-slot alternatives, the network performance measures they infer from various data sets proved useful to the models.

Of course, all studies discussed above use discrete methods to model departure time choice. In contrast, there has been a variety of relatively recent exploration of continuous methods using Cox's (1972) proportional hazards (PH) specification. For example, Wang (1996), Bhat (1996), Bhat and Steed (2002) and Lee and Timmermans (2007) used hazard-based models to model departure time on a continuous scale. Essentially, the ending time of the activity prior to the trip start is viewed as a "failure" or death of the activity.

There have been several applications of such continuous failure time models to different kinds of events/activities like onset of disease, equipment failures, job terminations, arrests and, more recently, departure time. Wang (1996) used a parametric-baseline hazard-rate model of durations, to estimate the revealed preferences of Canadians in activity start times. The estimates were then applied in a scheduling program, in order to examine how trip makers maximize their "total timing utility", and simulation results were used to quantify the tradeoffs between travel time and scheduling choices. Finally, a commuter-based network equilibrium was established on this basis. Bhat (1996) estimated a hazard-based model from grouped (7.5-minute interval-level) data of shopping durations (during work-to-home travel). He used a nonparametric baseline hazard distribution while parametrically controlling for covariate effects. Empirical results indicated significant effects of unobserved heterogeneity on shopping duration. Bhat and Steed (2002) used a similar semi-parametric model, to estimate departure times for shopping trips in Dallas-Fort Worth household survey data. Their continuous-time model was used to forecast temporal shifts in shopping trips, in order to evaluate the effects of congestion pricing.

Heckman and Singer (1984) illustrated means of incorporating heterogeneity via a non-parametric, finite-mixture specification. Such specification requires no prior assumptions, thereby minimizing the impacts of distributional assumptions that are standard in Cox's (1972) classic model. Most recently, Lee and Timmermans (2007) used a latent class

specification for the accelerated hazard model to capture heterogeneity and propensity to accelerate (or decelerate) activity durations. Theirs is the first example of a latent class accelerated hazard (LCAH) model for activity diary data on activity types (three out-of-home and two in-home). The data were collected in Eindhoven, The Netherlands, and the results suggest that heterogeneity relates strongly to demographic variables.

There also has been some exploration of a combination of discrete and continuous models. Mannering et al. (1990) employed such an approach to model route choices and departure times during morning work trips. They estimated a logit-type model for route choice and a continuous model of departure times. These were incorporated into a dynamic equilibrium framework, and several simulations were performed. Hamed and Mannering (1993) used a discrete/continuous framework for activity participation decisions, and associated travel times and activity durations. They modeled the "home-stay" duration using a Weibull survival model.

All studies mentioned above analyze departure time preferences either for work or non-work trips. There is almost no literature that examines the distribution of trips throughout the day, across trip purposes, on a continuous time scale. In order to address the gap, this paper looks into predicting continuous departure times for travelers and trips of various types. Most prior studies use a proportional hazards (PH) model (Cox, 1972), or some variation of it, to model departure time on a continuous scale. But as Kalbfleish and Prentice (1980) point out, the PH model's semi-parametric specification does not postulate a direct relationship between the covariates and the failure time without any restriction on the baseline hazard. In contrast, accelerated failure time (AFT) models use a log-linear specification and allow for straightforward interpretation of the covariate effects. Lee and Timmermans (2007) also point out such benefits.

Parametric AFT methods permit closed-form expressions for tail area probabilities, and their results are easier to interpret than those of semi or non-parametric models. The lognormal model is a pure version of AFT models, while the "Weibull model is a special case of both PH and AFT models" (Kalbfleish and Prentice, 1980, p. 34). For this reason of flexible interpretation, the Weibull is typically used. An extension of the Weibull model that includes unobserved heterogeneity via a gamma distribution has been very popular in the frequentist literature. However, in spite of promising research results by several authors, and as noted by Anderson et al. (1993), formal and completely satisfactory justifications of these likelihood-based methods await more results, related to their asymptotic and convergence properties. Furthermore, as Ibrahim et al. (2001, p.102) note, "none of these likelihood-based methods directly maximizes the full likelihood given the data, and the small sample properties of these estimators have yet to be studied. Thus, Bayesian methods are attractive for these models since they easily allow analysis using the full likelihood, and inference does not rely on asymptotics." (Ibrahim et al., 2001. pg). This paper uses parametric AFT models.

In view of these issues, this paper estimates and examines parametric (rather than semi-parametric) AFT models within a Bayesian framework, in order to model unimodal data. It then compares these results with those based on various other distributional assumptions, like the lognormal and Weibull (with and without latent heterogeneity). No previously published studies have modeled multimodal data, though this feature is certainly very common in aggregate departure time data (for certain trip purposes) and likely exists, to some extent, at the level of individuals (e.g., for business meetings, which generally may not take place during the noon hour). This paper develops a normal mixture model specification for such

data and illustrates those results for HBNW trip data. Of course, a key advantage of the Bayesian approach is that it allows for complicated hierarchical model specifications while facilitating estimation of predictive distributions; in contrast, maximum likelihood estimation methods cannot do so (see, e.g., Ibrahim et al. 2001).

To this end, this paper develops a Bayesian estimation of several parametric duration models for trip departure time with and without control for unobserved heterogeneity for unimodal data and also develops mixture models for multimodal data. These models are estimated and evaluated using data from the 1996 Austin Travel Survey (ATS).

3. Model description

Data that represent an interval of time are called duration data. Departure times (generally) mark the end of an activity's duration. Our interest lies in understanding how the covariates of an individual affect the distribution of these times. This requires assuming a base start time to compute the duration of activities that occur prior to one's departing for a trip. The choice of this time point "origin" is often problematic in continuous time models (Allison 1995), leading to somewhat different results for different origin choices. In ambiguous cases, this is dealt with by choosing the most intuitive origin. Here, the origin has been chosen as midnight, i.e., all the durations are calculated with respect to midnight. For example, an 8 am departure time implies an 8-hour (480-minute) event duration.

Parametric duration models using distributions like the exponential, lognormal and Weibull have been popular in the time-to-event/time-to-failure and survival analysis literature (see, e.g., Wang (1996)). This section describes the parametric AFT formulation of such models using the lognormal and Weibull distributions (as described in Kalbfleisch and Prentice (1980)). In addition, the Weibull-based model is extended to include unobserved heterogeneity among individuals, and a normal mixture specification is described for modeling bi-modal behaviors.

3.1 Accelerated failure time specification

Let the natural log of duration t (of an activity that occurs before taking a trip) be related to the covariates through a linear model,

$$\log(t) = \beta X + \sigma w \tag{1}$$

where, X is the set of covariates, β is a vector of parameters to be estimated, σ is the standard deviation of the transformed t value, and w is a random variable having a specified distribution.

Eq (1) is the most generic form of an AFT model. It not only provides for an easier interpretation in terms of $log(t)$ but also specifies that the effect of the co-variable as multiplicative on t rather than on the hazard function (as in the case of PH model). Different AFT models can be developed by varying the distribution of w. For example, if w is assumed to be a standard normal, it leads to a lognormal AFT model; and, if w is assumed to have an extreme value distribution, it results in a Weibull AFT model. These are described below.

3.1.1 Lognormal model

When w is assumed to have a standard normal distribution the density of t is log-normal, and is given by,

$$f(t|\mu,\tau) = \sqrt{\frac{\tau}{2\pi}}\frac{1}{t}\exp\left(-\frac{\tau}{2}(\log t - \mu)^2\right); t > 0 \tag{2}$$

where $\mu = \beta X$ and $\tau = 1/\sigma^2$, from Eq (1).

If T is defined as the duration during which an individual i does not take a trip (as measured relative to some base starting time, like midnight of the evening before), then the distribution $F(t)$ provides the probability that the duration for which an individual does not take a trip is less than a certain duration t. This is the same as the probability of taking a trip before a duration t elapses. The associated density $f(t)$ then offers the trip's departure time probabilities over the course of the day Moreover, the probability that a person takes a trip between times t_1 and t_2 can be obtained by integrating $f(t)$ from t_1 to t_2. Under this definition the likelihood of a trip starting at time t (relative to the base start time) for an individual i can be written as follows:

$$f(t|X_i,\beta,\tau) = \sqrt{\frac{\tau}{2\pi}}\frac{1}{t}\exp\left(-\frac{\tau}{2}(\log t - X_i\beta)^2\right) \tag{3}$$

Assuming, all reported trips (across respondents) are independent of one another, the joint likelihood for n trips is:

$$l(t|xi,\beta,\tau) = \prod_{i=1}^{n}\sqrt{\frac{\tau}{2\pi}}\frac{1}{t}\exp\left(-\frac{\tau}{2}(\log t - X_i\beta)^2\right) \tag{4}$$

3.1.2 Weibull model

If we let w have an extreme value distribution, the duration T before a respondent's trip will have a Weibull density (Kalbfleisch and Prentice [1980], pg 32). Letting the parameters of the Weibull density be α and λ, we have the following:

$$f(t|\lambda,\alpha) = \alpha t^{\alpha-1}\lambda \exp(\lambda t^\alpha) \tag{5}$$

where $\alpha = 1/\sigma$, $\lambda = \exp(X\beta^*)$ (thereby ensuring non-negativity of the hazard rate) and $\beta^* = -\beta/\sigma$. X is a set of covariates and β^* is a vector of parameters to be estimated. X, β and σ are as in Equation (1).

It follows that the likelihood of a trip starting at time t from the designated starting time (taken to be midnight here) for an individual i can be written as follows:

$$f(t|X_i,\alpha,\beta^*) = \alpha t^{\alpha-1}\exp(X_i\beta^*)\exp(-t^\alpha \exp(X_i\beta^*)) \tag{6}$$

The choice of designated starting time can be an important consideration. In most cases, one is modeling the duration of an activity or the time since completion of one's last activity, so

the starting time is obvious and implicit. Here, however, departure times (based on the clock) are being modeled, and the origin time was chosen to be midnight, for a couple reasons. First, lognormal and Weibull distributions are unimodal, so starting time must far from the data's mode. Secondly, a midnight origin facilitates interpretation of model results (since no shifting needed). A 3 am start time also was investigated here (since that is the cut-off time for many one-day travel diaries), and the results were nearly identical to those shown here. Assuming that all trips in one's data set are independent of one another (which may not be the case for members of the same household), the joint likelihood for n trips is:

$$\ell(t|x,\alpha,\beta^*) = \prod_{i=1}^{n} \alpha t^{\alpha-1} \exp(X_i\beta^*)\exp(-t^\alpha \exp(X_i\beta^*)) \tag{7}$$

The independence of trips assumption is largely a reasonable assumption here, since different trip types were modeled separately; so so almost no one had two trips of the same trip type. However, in the model of HBNW trips many travelers were headed out and back during the survey day. Ideally, some sort of correlation in departure times would be permitted, and Bayesian methods would be a natural way to accommodate such correlation.

3.2 Weibull model with unobserved heterogeneity (UH)

One way to accommodate UH in a Weibull model is via a multiplicative error term v_i in each respondent's hazard function, as follows:

$$h(t, X_i) = v_i \alpha t^{\alpha-1} \exp(X_i\beta^*) \tag{8}$$

The associated density and likelihood functions are as followsKalbfleish and Prentice (1980)):

$$f(t|X_i,\alpha,\beta) = v_i \alpha t^{\alpha-1} \exp(X_i\beta^*)\exp(-t^\alpha v_i \exp(X_i\beta^*)) \tag{9}$$

$$l(t|x,\alpha,\beta) = \prod_{i=1}^{n} v_i \alpha t^{\alpha-1} \exp(X_i\beta^*)\exp(-t^\alpha v_i \exp(X_i\beta^*)) \tag{10}$$

To ensure non-negativity, v_i is assumed to be gamma distributed here, with a mean of 1 and an unknown variance δ (to be estimated).

3.3 Normal mixture model

A mixture of normals provides a relatively flexible model for estimation of densities in a Bayesian framework (Roeder and Wasserman, 1997). Such models help capture any multimodality existing in the data. Consider, for example, a set of observations y_1, $y_2,...,y_j,...,y_n$ to be modeled through a normal mixture distribution (as a mixture of k normals):

$$f(y_j) = \sum_{i=1}^{k} p_i N(\mu_i, \sigma_i^2) \tag{11}$$

where p_i is the probability of individual j's behavior belonging to group i (of the k groups), so that $\sum_{i=1}^{k} p_i = 1$, and $N(\ )$ is the normal density function.

Here, we assume a lognormal density for duration t, which makes the distribution of $y = \ln(t)$ normal. Moreover, the (aggregate) home-based non-work (HBNW) trip data appear to be largely bimodal in nature, so we examine a mixture of two normal distributions, in order that each observation y_j is presumed to come from one of two groups. The model can be formulated as follows:

$$f(y|\mu_1, \mu_2, \sigma_1{}^2, \sigma_2{}^2) = pN(\mu_1, \sigma_1{}^2) + (1-p)N(\mu_2, \sigma_2{}^2) \tag{12}$$

Assuming $\mu_1 = \beta_1 X_i$, $\mu_2 = \beta_2 X_i$, $\tau_1 = 1/\sigma^2{}_1$, and $\tau_2 = 1/\sigma^2{}_2$, then:

$$f(y|X_i, \beta_1, \beta_2, \tau_1, \tau_2) = p\sqrt{\frac{\tau_1}{2\pi}}\exp\left(-\frac{\tau_1}{2}(y-X_i\beta_1)^2\right) + (1-p)\sqrt{\frac{\tau_2}{2\pi}}\exp\left(-\frac{\tau_2}{2}(y-X_i\beta_2)^2\right) \tag{13}$$

Thus, the likelihood for n (independent) departures will be,

$$f(y|X_i, \beta_1, \beta_2, \sigma_1, \sigma_2, p) = \prod_{i=1}^{n}\left(p\sqrt{\frac{\tau_1}{2\pi}}\exp\left(\frac{-\tau_1(y-X_i\beta_1)^2}{2}\right) + (1-p)\sqrt{\frac{\tau_2}{2\pi}}\exp\left(\frac{-\tau_2(y-X_i\beta_2)^2}{2}\right)\right) \tag{14}$$

4. Model Estimation

A parametric Weibull model implies that the departure time distribution (for any individual) is unimodal. Of course, HBW trips are defined to include both directions of travel, and thus generally exhibit two peaks each day. In order to more accurately represent these distinct behaviors, HBW trips were separated into home-to-work trips and work–to-home trips, and separate models were estimated. NHB trips also appeared unimodal (in the aggregate) and thus were modeled similarly. In contrast, HBNW trips appear multimodal and a to/from separation of peaks will not resolve this multimodality. Thus, a mixture model formulation was used for HBNW trips. All parameters are estimated here using a Bayesian approach. (Gelman et al. 2003)

The posterior distributions are obtained as a product of the prior distributions and the data's likelihood functions. In canonical notation, the four model specifications are as follows:

Lognormal model:

$$f(\beta, \sigma|t, x_i) \propto l(t|x_i, \beta, \sigma) \cdot f(\beta) \cdot f(\sigma) \tag{15}$$

Weibull model:

$$f(\alpha, \beta, \delta|t, x_i) \propto l(t|\alpha, \beta, x_i) \times f(\alpha) \times f(\beta) \times f(\delta) \tag{16}$$

Weibull model with UH:

$$f(\alpha, \beta, \delta | t, x_i) \propto l(t | \alpha, \beta, x_i) \times f(\alpha) \times f(\beta) \times f(\delta) \qquad (17)$$

Mixture of two normals:

$$f(\beta_1, \beta_2, \tau_1, \tau_2, p | t, x_i) \propto l(t | x_i, \beta_1, \beta_2, \tau_1, \tau_2) \cdot f(\beta_1) \cdot f(\beta_2) \cdot f(\tau_1) \cdot f(\tau_2) \cdot f(p) \qquad (18)$$

where $f(\tau)$, $f(\alpha)$, $f(\beta)$ and $f(\delta)$ are the prior densities of parameters τ, α, β and δ, respectively.

The prior distributions for τ, α and δ are assumed to be independent gamma distributed (in order to ensure non-negativity), and the β's are assumed to follow a multivariate normal. The prior for p is a Dirichlet distribution (to ensure that the probabilities sum to 1), and this reduces to a beta distribution for a mixture of just two normals. All priors are non-informative with appropriate mean values (e.g., zero means for all beta values and unit means for gamma size and shape parameters) and large variances (since there is little to no information on these values prior to the model's estimation).

The joint posterior distributions do not have a closed form, so Markov chain Monte Carlo (MCMC) simulations were developed using WinBUGS1.4 (or BUGS [http://www.mrc-bsu.cam.ac.uk/bugs/]) software, which simulates posterior distributions for known likelihood specifications using standard priors. BUGS (Bayesian inference Using Gibbs Sampling) is an open-source software for the Bayesian analysis of complex statistical models using MCMC methods. The basic idea is to successively sample from the conditional distribution of each parameter given all the other parameters, and this process eventually provides samples from the joint posterior distribution of the unknown random quantities. Results for the first three model specifications (each with three applications, for home-to-work, work-to-home, and NHB trips) and one application of the mixture of normals (for HBNW trips) are all based on 10,000 simulations, following a 1000-draw burn in. In the interest of brevity, convergence monitoring is not shown in this paper. Here we merely note that standard procedures described in BUGS were used to ensure proper convergence.

Model selection was based on stepwise deletion of variables based on statistical and practical significance of the associated coefficients. Coefficients with t-statistics less than 1 generally were removed; however, several that appeared intuitive were left in the specification, even if only marginally significant (in a statistical sense). And some coefficients which were statistically significant in one specification were left in other specifications (though insignificant) to allow for comparison.

Model evaluation was performed by comparing the model-predicted departure time densities to the actual histogram of departure times, and model-predicted departure densities were estimated as the average of all individuals' departure time densities. The process is not so trivial for the Weibull with UH model, due to conditioning on v. Details of this process are provided in the Appendix to this paper.

5. Empirical Analysis

The analysis was carried out using a total of 3,189 HBW trips (of which 1,717 are home-to-work trips and 1,472 are work-to-home trips), 4,856 NHB trips, and 7,325 HBNW trips. Tables 1 and 2 provide descriptive statistics for all variables used. Tables 3, 4, and 5 present the estimation results for the lognormal, standard Weibull and Weibull with UH models for the home-to-work, work-to-home and NHB trip types, respectively. Table 6 shows the results

of the normal mixture model for HBNW trips. Figure 1's four images illustrate model fit for each of these models against actual data information.

5.1 Interpretation of Coefficients and Marginal Effects

As described in the specification, μ is parameterized as βX for the lognormal model. Thus, the signs of β values indicate whether an individual with associated attribute X leave later or earlier in the day. For example, if the sign on the coefficient of age is negative, it indicates that older individuals are more likely to depart earlier for that trip type (essentially experiencing a shorter activity duration prior to their departure). The coefficients corresponding to each peak of the bi-modal model also can be interpreted similar to the case of the lognormal model, but relative to the associated peak. For example, if the sign on the coefficient of student is positive for the AM peak, it indicates that the students are more likely to depart later in the AM period for that trip purpose. As noted earlier, in both the Weibull models, $\beta^* = -\beta / \sigma$. Thus, an appropriate interpretation of a positive beta in these two models is that people belonging to the corresponding demographic class will depart sooner (thanks to a higher failure rate), thus ending their prior activity earlier. This is in contrast to interpretation of signs for the lognormal model, discussed above. For example, if the coefficient estimate for the male variable is positive in the Weibull models, it would be correct to say that men take such a trip earlier than women, everything else held constant. (The opposite would be true for the lognormal model's interpretation.)

Based on these tendencies, results for each control variable shown in Tables 3, 4 and 5 are consistent across all three models. In other words, if the coefficient on a control variable is positive in the lognormal model, it is negative in both of the Weibull models and vice-versa. Of course, as expected, the magnitude and statistical significance of the coefficients vary.

The percentage increase/decrease in expected departure time for a unit change in X can be computed as $100(e^{\beta} - 1)$ and $100(e^{-\beta/\alpha} - 1)$ for the lognormal and Weibull models, respectively. For indicator variables, these equations give the percentage increase (lognormal model) or decrease (Weibull models) in the expected departure time for those exhibiting such characteristics ($X=1$), relative to others ($X=0$) – and relative to the midnight start time. These values are presented (parenthetically) in Tables 3, 4, and 5, and the key results (in terms of minutes, rather than percentages) are discussed here.

A variety of demographic and trip-related variables were found to be highly statistically significant in all estimated models. In the ***home-to-work*** models, the attributes of greatest practical significance (in terms of their effect on departure times) were found to be (in order of importance, and using the Weibull model): (a) Hispanic ethnicity (resulting in a 10% or 40-minute earlier departure time than the base ethnicity of Asian, and a 20-minute earlier departure time than Caucasians), (b) employment status (with full-time employees departing 35 minutes earlier than part-time employees), (c) trip attributes (external trips leaving 27 minutes later after controlling for distance), and (d) flexible work hours (25 minutes later than those without flexible work hours).

In the Weibull model of ***work-to-home*** departure times, the most practically relevant attributes were found to be: (a) ethnicity (Hispanics generally departing 20 minutes later than Asians and at the same time as Caucasians), (b) external-zone destination (which occur 115 minutes later than intraregional trips, after controlling for trip cost), (c) day of the week (with workers departing an average of 20 minutes earlier on Fridays), and (d) mode (with

carpoolers and solo drivers departing an hour later than bus users). Under both the Weibull models, the key variables were somewhat fewer: (a) mode, (b) external trips, (c) ethnicity, and (d) job type.

While the model results suggest that individuals tend to depart earlier *from* work on Fridays, no day-of-week effects were found in home-to-work trips (i.e., departures from home). However, there is some evidence here that individuals leave later for work in the summertime (though there is no such effect for work-to-home trips). Interestingly, the presence of children in the household appears to have little effect on departure time choice in home-to-work trips, but an earlier departure time for parents (or those living in households with children) is evident in the work-to-home trips. Combining the effects from the two HBW models, one can conclude that individuals with higher incomes, Caucasians and Hispanics, full-time workers, and those commuting solo tend to work the longest days. In contrast, those enjoying flexible work schedules and those working at offices tend to experience more compressed work days (though their lunch breaks may also be shorter).

Table 5's **NHB trip** model results indicate that a work-trip purpose, the presence of children, traveler age and employment status are the most statistically and practically significant variables. Most results are intuitive: for example, males, those in households with children present and those making a NHB trip to or from work tend to make such trips earlier in the day (15 minutes earlier for males [versus females], 45 minutes in presence of children, and 80 minutes when for work purposes, as computed using the Weibull model's results). NHB trips also appear to occur earlier in summer (50 minutes earlier than in the fall and wintertime) but somewhat later in the springtime (by 18 minutes) and on Fridays (by 24 minutes).

Table 6 presents the bi-modal model results for **HBNW** trips. These indicate that the most important determinants of departure time for such trips are the presence of children (which tend to occur 30 minutes earlier in the PM period, when using the normal mixture model), employment status (with full-time workers leaving 50 minutes earlier [in the AM period] than their part-time counterparts), mode of travel (with bus users departing 45 minutes earlier than car drivers, in the AM period) and travel cost (with every added dollar of cost associated with a 5 minute later AM departure time and 5 minute earlier PM time). The directional effects of each of the variables and their comparison for AM and PM peaks is accomplished by focusing on coefficient signs. For example, retired individuals tend to take HBNW trips later in the AM period (β_{AM} = +0.0574) but earlier in the PM period (β_{PM} = -0.0673). Interestingly, such trips are taken later in the AM period in spring and later in the PM during summer.

Clearly, there is strong consistency across model results by trip type, and many explanatory variables are associated with very practically significant effects. Nevertheless, it is important to recognize which of the model specifications performs best, so that analysts can focus on their results. The following section describes such evaluations.

5.2 Model Comparisons

Models can be compared in multiple ways. Two of the most common are the magnitude of factor impacts, as well as goodness of fit. Both of these evaluation approaches are examined here.

5.2.1 Coefficients and Marginal Effects

For home-to-work trips most parameter estimates were similar (in terms of having the same general effect on departure time) in all three models. However, the presence of children was found to be both statistically and practically significant in the Weibull model with UH, but not in the other two models. Summertime and educational employment variables were found to be statistically insignificant in the Weibull model with UH, but significant in the other two models. Overall, the results from the Weibull model appear somewhat more intuitive. In terms of marginal impacts of control variables, the lognormal and Weibull models offered higher values than those of the Weibull model with UH.

For work-to-home trips the coefficients were similar across models, except that income and office-employment variables were statistically significant only in the lognormal model. In contrast, African-American race and retail occupation variables were statistically significant only in the two Weibull models. The coefficients for both Weibull models are roughly the same. This is consistent with the estimate of v_i's standard deviation being quite small (0.05) – and only marginally significant (since v_i averages to 1). Essentially, unobserved heterogeneity in the work-to-home trips model is slight, resulting in more similar outcomes. In terms of factor magnitudes, the lognormal model exhibits stronger responses to variations in control variables than do either of the Weibull models.

For NHB trips the results are consistent across all three models except for the cost variable, which was statistically insignificant in lognormal model , and the summertime variable, which was found to be statistically (and practically) significant only in the lognormal model. In terms of marginal impacts, all factors offered values in the same range across all three models.

5.2.2 Goodness of fit

Spiegelhalter et al.'s (2002) Deviance Information Criterion (DIC) was used to compare the models' overall performance. The common Akaike and Bayesian Information Criteria (AIC and BIC) are limiting cases of the DIC, which is composed of two terms: one accounts for model fit and the other for model size. Models with smaller values of DIC are preferred. (This is also the rule for model choice under the AIC and BIC, which are felt to be inferior metrics. [Spielgelhalter et al (2002), pp 185].)

DIC values illustrate how the Weibull with UH is preferred for home-to-work trips (suggesting that latent heterogeneity is important for this trip type), while the standard Weibull is only slightly preferred to the Weibull with UH for work-to-home trips (DIC = 19,668 versus 19,682) (an indication that less latent variation exists for this trip type). Consistent with Table 3's results, Figure 1(a) illustrates population density function estimates for trip departure times across the sample population, supporting the conclusion that the Weibull with UH substantially outperforms the lognormal in both cases and the standard Weibull in the home-to-work case. Figure 1(b) clearly illustrates how both Weibull model predictions for work-to-home trips are nearly the same (consistent with their nearly equal DIC values). Unfortunately, while both out-perform the lognormal model, neither performs impressively in the aggregate comparisons. Both exhibit far too much spread compared to the observed distribution of work-to-home travel times.

A comparison of Figures 1(a) and 1(b) raises an important question: Why is it that work-to-home trips are so much harder to predict than home-to-work trips with these models? Perhaps the symmetry and slightly leftward skew of the work-to-home departure time pattern

is a problem, in addition to the rather clear bi-modality? Perhaps business practices govern departure times from work, while home location, presence of children, and commute times play far more important roles in the home-to-work decision? Such a scenario is easy to imagine and suggests that other control variables may be key for the work-to-home model.

For NHB trips, the DIC values suggest that the Weibull is the best model, as evident in Figure 1 (c). The Weibull model also performs best in replicating the aggregate data. Moreover, in terms of unobserved heterogeneity, it was found to be only marginally significant statistically, with no real practical significance. Thus, there seems to be little difference in Weibull model results with and without UH, suggesting the attributes already controlled for in the model capture the NHB trip departure time behavior as well as they can.

The bi-modal normal mixture model for HBNW trips predicts the distribution's two modes quite well (Figure 1 (d)), but missed the AM peak's spread in departure times quite a lot. This may be due to the abrupt nature of the AM peaking (i.e, a very large number of trips in a 30-minute interval, which is difficult to capture in such models).

6. Conclusions

The formulation, estimation, and evaluation of Bayesian parametric duration models using different distributional assumptions were demonstrated here, using models for various trip purposes from the 1996 Austin Travel Survey. The departure time models for home-to-work and work-to-home trip purposes were estimated using lognormal, standard Weibull and Weibull with unobserved heterogeneity (UH) models. These specifications were compared using a relatively new goodness of fit statistic, called the DIC. The Weibull model with UH was found to perform best for both trip purposes, both in terms of goodness of fit and predictive (aggregate) distributions for HBW trip departure times. However, recognition of unobserved heterogeneity in the home-to-work departure time model improved model fit much more than in the work-to-home application.

The Weibull model performed best in fitting the NHB trip data, and the normal mixture model did well in terms of predicting the modes of the HBNW departure time distribution. The empirical analyses provided intuitively appealing results, allowing for behavioral understanding of departure time patterns across user groups, as well as generalized travel costs, recognizing both time and money. The Bayesian approaches employed here provide a valuable alternative to the presently MLE-based methods, in terms of the breadth of information obtained on parameter distributions, unbiased comparison methods for different (non-nested) models (via the DIC), and the availability and user-friendliness of the associated software (WinBUGS). Gadda (2006) performed a comparison of the Bayesian and MLE methods for the Weibull specification, and both sets of results were very similar for this data set. Bayesian methods should be of even greater use in future analyses, when some prior information exists for inclusion, which is not possible with MLE-based methods.

In terms of immediate, practical applications of this work's results, estimates explain how departure time choices are likely to vary according to traveler attributes. Such information is likely to be useful in anticipating the peak travel times for different cities and different areas of a network, as well as the travelers most impacted by new policies. Key features of time-of-day dependent policies include variable-rate toll schedules, carpool lane hours of operation, and staggered work start time decisions.

While the results shared here appear very promising, some limitations remain. For example, departure time models in this application require a base start time, which affects results slightly. Moreover, standard approaches do not have an end time, when, in fact, 24 hours is the effective constraint used here, impacting left-skewed distributions (e.g., late-night activities like movie watching). This constraint would not apply, of course, if one wanted to use these models for activity durations, as individuals could remain in an activity indefinitely. Other limitations in the current models, as applied here, include the independence of trips assumption which is mostly handled by "divorcing" outbound and return trips (which may be intimately related in reality) and an inability to control for time-of-day-varying travel costs, which are important for various applications (such as variable tolling). Also, the use of parametric distributions is less flexible than semi-parametric and non-parametric approaches and may lead to erroneous conclusions if the assumed underlying distributions are incorrect.

The next step for enhancing the models presented here is to allow for time-varying covariates (as used in Popuri et al. 2008 and Abou-Zeid et al. 2006, for example) and simultaneous analysis of inter-activity durations, in order to resolve most of the issues highlighted here. This can be done by studying all trips made by an individual in a day and estimating durations of all activities, and then re-formulating the model in terms of these durations. Also, recent research in Bayesian nonparametric modeling shows that these models typically tend to outperform parametric models in both fit and prediction. Among other reasons, nonparametric priors allow for high levels of skew and kurtosis in the data that parametric models typically fail to capture (see, e.g., Dey and Rao's 2006 compendium of papers). Past, current, and future research will no doubt evaluate the value of these models in transportation research. In particular, developing hierarchical semi-parametric scale mixture models, which better account for uncertainty in the observed data, is one area of research that should be pursued. In addition, a comparison of this paper's methodology with those applied in prior works using hazard-based models of departure time is of interest.

The methodologies presented in this paper should assist the development of more robust models for continuous departure time choices like incorporating these in the tour/activity based travel data and modeling systems. Also, the outputs of theses modles could be useful inputs to emerging dynamic traffic assignment models and traffic control measure evaluation (such as implementation of congestion pricing policies). Departure time choice is a complex phenomenon, and continuous-time models are needed, to capitalize on the power of emerging dynamic traffic assignment models while providing the requisite data.

Acknowledgements

Appendix

As noted, model evaluation was performed by comparing model-predicted departure time densities to the actual histograms of departure times. These model-predicted departure densities were estimated as the average of all individuals' departure time densities (evaluated at the average values of parameter estimates) in the data set. Evaluating these densities for the lognormal, Weibull and normal mixture models was a straightforward process. However, to find these densities for the Weibull model with UH, the conditioning on v_i had to be removed. This resulted in the following departure time density:

$$f(t|X_i, \alpha, \beta) = \frac{\alpha t^{\alpha-1} \exp(X_i \beta)}{\left(\dfrac{t^{\alpha} \exp(X_i \beta)}{\delta} + 1\right)^{\delta+1}} \tag{A1}$$

The derivation of Eq. (A1) is as follows: Individual i's conditional departure time is given by:

$$f(t|X_i, \alpha, \beta, v_i) = v_i \alpha t^{\alpha-1} \exp(X_i \beta) \exp(-t^{\alpha} v_i \exp(X_i \beta)) \tag{A2}$$

Now,

$$f(t|X_i, \alpha, \beta) = \int_0^\infty v_i \alpha t^{\alpha-1} \exp(X_i \beta) \exp(-t^{\alpha} v_i \exp(X_i \beta)) f(v_i) dv_i \tag{A3}$$

Letting $\lambda_i = \exp(X_i \beta)$,

$$f(t|\alpha, \lambda_i) = \int_0^\infty v_i \alpha t^{\alpha-1} \lambda_i \exp(-t^{\alpha} v_i \lambda_i) f(v_i) dv_i \tag{A4}$$

$$\Rightarrow f(t|\alpha, \lambda_i) = \alpha t^{\alpha-1} \lambda_i \int_0^\infty v_i \exp(-t^{\alpha} v_i \lambda_i) f(v_i) dv_i$$
(A5)

Since v_i is gamma distributed with mean 1 and variance $1/\delta$, the following holds:

$$f(v_i) = \frac{\delta^\delta}{\Gamma(\delta)} v_i^{\delta-1} \exp(-v_i \delta). \tag{A6}$$

Substituting for $f(v_i)$ in the integrand:

$$\Rightarrow f(t|\alpha, \lambda_i) = \frac{\delta^\delta}{\Gamma(\delta)} \alpha t^{\alpha-1} \lambda_i \int_0^\infty v_i \exp(-t^{\alpha} v_i \lambda) v_i^{\delta-1} \exp(-v_i \delta) dv_i \tag{A7}$$

$$= \frac{\delta^\delta}{\Gamma(\delta)} \alpha t^{\alpha-1} \lambda_i \int_0^\infty v_i^{\delta} \exp(-(t^{\alpha} \lambda + \delta) v_i) dv_i$$

Substituting $y = (t^{\alpha} \lambda_i + \delta) v_i$, one has the following:

$$= \frac{\delta^\delta}{\Gamma(\delta)} \alpha t^{\alpha-1} \lambda_i \int_0^\infty \frac{y^{\delta}}{(t^{\alpha} \lambda + \delta)^{\delta}} \exp(-y) \frac{dy}{(t^{\alpha} \lambda + \delta)}$$

$$= \frac{\delta^\delta \alpha t^{\alpha-1} \lambda_i}{\Gamma(\delta)(t^{\alpha} \lambda + \delta)^{\delta+1}} \int_0^\infty y^{\delta} \exp(-y) dy$$

15

$$= \frac{\delta^{\delta} \alpha t^{\alpha-1} \lambda_i}{\Gamma(\delta)(t^{\alpha}\lambda_i + \delta)^{\delta+1}} \Gamma(\delta+1)$$

$$= \frac{\delta^{\delta+1} \alpha t^{\alpha-1} \lambda_i}{(t^{\alpha}\lambda_i + \delta)^{\delta+1}}$$

Therefore, $f(t|\alpha, \lambda_i) = \dfrac{\alpha t^{\alpha-1} \lambda_i}{\left(\dfrac{t^{\alpha}\lambda_i}{\delta} + 1\right)^{\delta+1}}$ (A8)

Substituting $\lambda_i = \exp(X_i\beta)$ into the last equation completes the derivation.

These results were used to calculate the departure time density of each individual, as well as their average, thus providing the population's predictive density for a comparison with observed distributions.

References

Abkowitz, Mark D. (1981) An Analysis of the Commuter Departure Time Decision. *Transportation,* Vol. 10, pp. 283-297.

Abou-Zeid, M., Rossi, T., and Gardner, B. (2006) Modeling Time-of-Day Choice in Context of Tour- and Activity-Based Models. *Transportation Research Record* No. 1981, pp. 42–49.

Allison, P.D. (1995) *Survival Analysis using the SAS system: A Practical Guide*, Cary, NC: SAS Institute Inc.

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

Backjin Lee , Harry J.P. Timmermans (2007) A latent class accelerated hazard model of activity episode durations. *Transportation Research* Part B, Vol. 41 , pp. 426-447.

Bhat, C.R. (1996) A Hazard-Based Duration Model of Shopping Activity with Nonparametric Baseline Specification and Nonparametric Control for Unobserved Heterogeneity. *Transportation Research* Part B, Vol. 30, pp. 189-207.

Bhat, C.R. (1998) An Analysis of Travel Mode and Departure Time Choice for Urban Shopping Trips, *Transportation Research* Part B, Vol. 32, No. 6, pp.361-371.

Bhat, C.R., and Steed. J.L. (2002) A Continuous-Time Model of Departure Time Choice for Urban Shopping Trips, *Transportation Research* Part B, Vol. 36, No. 3, pp. 207-224.

Ben-Akiva, M. E., and Lerman, S. R. (1985). *Discrete Choice Analysis: Theory and Application to Travel Demand*, MIT Press, Cambridge, Massachusetts.

Chin, Anthony T. H. (1990) Influences on Commuter Trip Departure Time Decisions in Singapore. *Transportation Research* Part A, Vol. 24, No. 5, pp. 321-333.

Cox, D. R. (1972) Regression models and life tables. *Journal of Royal Statistical Society*, Series B, Vol. 34, pp.187–220.

Dey, D., and Rao, C.R., (2006), *Handbook of Statistics*, Volume 25, Elsevier, North Holland.

Gadda, S. (2006) 'Ranked Logit Analysis, Continuous Models of Departure Time, and Quantifying Uncertainty in AADT Estimates', Masters Thesis, The University of Texas at Austin.

Gelman, A., Carlin, J. B., Stern H. S., and Rubin, D. B. (2003) *Bayesian Data Analysis,* 2nd ed, Chapman & Hall/CRC

Hendrickson, Chris, and Plank, E. (1984) The Flexibility of Departure Times for Work Trips. *Transportation Research* A, Vol. 18, No. 1, pp. 25-36.

Ibrahim, J.G., Chen, M.-H., Sinha, D. (2001) *Bayesian Survival Analysis*, New York: Springer-Verlag.

Kalbfleisch, J.D., and Prentice, R.L. (1980). *The Statistical Analysis of Failure Time Data*, Wiley, NY.

Mannering, F., Abu-Eisheh, S., Arnadottir, A. (1990) Dynamic traffic equilibrium with discrete/continuous econometric models. *Transportation Science* 24(2), pp. 105-116.

Hamed, M., Mannering, F. (1993) Modeling travelers' post-work activity involvement: toward a new methodology. Transportation Science 27(4), pp. 381-394.

Heckman, J., Singer, B (1984) A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica* 52, pp. 909–938.

McCafferty, Desmond, and Fred L. H. (1982) The Use of Multinomial Logit Analysis to Model the Choice of Time to Travel. *Economic Geography*, Vol. 36, No. 3, pp. 236-246.

McFadden, D. (1978), 'Modeling the choice of residential location', in A. Karlqvist, L. Lundqvist, F. Snickars, and J. Weibull, eds., *Spatial Interaction Theory and Planning Models*, North-Holland, Amsterdam, pp. 75–96.

Popuri, Y., Ben-Akiva, M., and Proussaloglou, K. (2008) Time of day modeling in a tour-based context: The Tel-Aviv experience. Proceedings of the 87th Annual Meeting of the Transportation Research Board. Washington, D.C.

Small, Kenneth A. (1982) The Scheduling of Consumer Activities: Work Trips. The *American Economic Review*, Vol. 72, No. 3, pp. 467-479.

Spiegelhalter D. J., Best N. G., Carlin B. P., and van der Linde A. (2002) Bayesian measures of model complexity and fit (with discussion). *Journal of Royal Statistical Society* B. 64, pp. 583-640.

Spiegelhalter, D., Thomas, A., Best, N., and Lunn, D. (2003) *WinBUGS User Manua*l, Version 1.4, http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/manual14.pdf (accessed May 20, 2006).

Steed, J., and Bhat, C.R. (2000) On Modeling the Departure Time Choice for Home-Based Social/Recreational and Shopping Trips. *Transportation Research Record*, Vol. 1706, pp. 152-159.

Wang J.X. (1996) Timing Utility of Daily Activities and Its Impact on Travel. *Transportation Research* A, vol. 30, Vo. 3, pp. 189-206.

List of Tables and Figure

Table 1. Descriptive Statistics of Sample Data for HBW Trips

Table 2. Descriptive Statistics of Sample Data for NHB and HBNW Trips

Table 3. Final Model Results for Home-to-work Trips

Table 4. Final Model Results for Work-to-home Trips

Table 5. Final Model Results for NHB Trips

Table 6. Normal Mixture Model Results for HBNW Trips

Figure 1. Comparison of Model Predictions for Different Trip Purposes

Table1. Descriptive Statistics of Sample Data for HBW Trips

Variable	Description of the Variable	Home-to-work trips				Work-to-home trips			
		Min	Max	Mean	S.D	Min	Max	Mean	S.D
Dep. Time	Departure time (minutes after midnight)	170	1380	520.80	183.44	5	1440	991.89	194.76
Age	Age of respondent (years)	5	81	39.08	12.29	5	99	38.85	12.71
Income	Household income (divided by $10,000)	0.25	17.5	5.41	3.34	0.25	17.5	5.35	3.28
Kids	Presence of children under the age of 5 (1 = yes)	0	1	0.15	0.35	0	1	0.15	0.36
Gender	(1 = male)	0	1	0.50	0.50	0	1	0.51	0.50
Student	Student full time (1 = yes)	0	1	0.13	0.33	0	1	0.13	0.34
AfAm	African-American (1 = yes)	0	1	0.05	0.23	0	1	0.81	0.39
Hispanic	Hispanic (1 = yes)	0	1	0.11	0.31	0	1	0.06	0.23
Caucasian	Caucasian (1 = yes)	0	1	0.78	0.42	0	1	0.11	0.31
Full time	Full-time worker (1 = yes)	0	1	0.83	0.38	0	1	0.78	0.42
Flex Work	Enjoy flexible work-hours (1 = yes)	0	1	0.41	0.49	0	1	0.42	0.49
Retail	Retail sector employee (1 = yes)	0	1	0.16	0.37	0	1	0.18	0.38
Industrial	Industrial sector employee? (1 = yes)	0	1	0.12	0.33	0	1	0.12	0.33
Education	Work for an educational institution (1 = yes)	0	1	0.06	0.24	0	1	0.07	0.25
Office	Works in an office (1 = yes)	0	1	0.21	0.41	0	1	0.20	0.40
Government	Works in a government agency (1 = yes)	0	1	0.12	0.33	0	1	0.12	0.33
Drive alone	Drive alone (on trip in question, 1 = yes)	0	1	0.86	0.35	0	1	0.80	0.40
Shared ride	Shared ride (on trip in question, 1 = yes)	0	1	0.10	0.30	0	1	0.16	0.37
External	Travel to an external zone on trip in question (1 = yes)	0	1	0.01	0.10	0	1	0.01	0.10
Cost	Generalized cost of trip ($, in 1996)	0.205	34.8	5.03	3.49	0.205	30.71	5.23	3.62
Friday	Trip took place on a Friday (1 = yes)	0	1	0.16	0.37	0	1	0.16	0.37
Summer	Trip took place during summertime (June, July, & August)	0	1	0.03	0.18	0	1	0.03	0.18
Number of observations		1717				1472			

Table 2. Descriptive Statistics of Sample Data for NHB and HBNW Trips

Variable	Description of the Variable	Non-home-based trips				Home-based non-work trips			
		Min	Max	Mean	S.D	Min	Max	Mean	S. D
Dep. Time	Departure time (Minutes after 12am)	10	1440	875.56	291.06	10	1440	843.61	224.30
Age	Age of respondent (years)	5	92	38.37	16.64	5	99	34.77	20.23
Income	Household income (divided by $10,000)	0.25	17.5	5.83	3.68	0.25	17.5	5.40	3.64
Kids	Presence of children under the age of 5 (1 = yes)	0	1	0.16	0.37	0	1	0.18	0.38
Male	Gender of the person (1 = male)	0	1	0.49	0.50	0	1	0.50	0.50
Hispanic	Hispanic (1 = yes)	0	1	0.08	0.28	0	1	0.12	0.32
Caucasian	Caucasian (1 = yes)	0	1	0.84	0.37	0	1	0.78	0.41
Full time	Full-time worker (1 = yes)	0	1	0.62	0.49	0	1	0.41	0.49
Part time	Part-time worker (1 = yes)	0	1	0.08	0.28	0	1	0.10	0.29
Retired	Retired from employment (1 = yes)	0	1	0.07	0.25	0	1	0.10	0.30
Student	Student full time (1 = yes)	0	1	0.23	0.42	0	1	0.37	0.48
Retail	Retail sector employee (1 = yes)	0	1	0.09	0.29	0	1	0.08	0.26
Industrial	Industrial sector employee (1 = yes)	0	1	0.09	0.28	0	1	0.05	0.23
Education	Work for an educational institution (1 = yes)	0	1	0.04	0.20	0	1	0.03	0.17
Office	Works in an Office (1=yes)	0	1	0.19	0.39	0	1	0.11	0.31
Government	Works in a Government agency (1=yes)	0	1	0.09	0.29	0	1	0.06	0.25
Drive alone	Drive alone (on trip in question, 1 = yes)	0	1	0.52	0.50	0	1	0.38	0.49
Shared ride	Shared Ride (on trip in question, 1 = yes)	0	1	0.41	0.49	0	1	0.49	0.50
Trip cost	Generalized cost of trip ($, in 1996)	.13	34.80	2.54	2.63	0.13	23.35	2.16	2.06
Friday	Trip took place on a Friday (1=yes)	0	1	0.20	0.40	0	1	0.18	0.38
Summer	Trip took place during Summer (1=yes)	0	1	0.03	0.18	0	1	0.04	0.20
Spring	Trip took place during Spring (1=yes)	0	1	0.87	0.34	0	1	0.83	0.38
Number of observations		4856				7325			

Table 3. Final Model Results for Home-to-Work Trips

Attribute	Variable	Lognormal		Weibull		Weibull with UH	
		Beta (ME)	t-stat	Beta (ME)	t-stat	Beta (ME)	t-stat
	Alpha(α)			2.997	58.638	7.884	6.116
	Constant(b0)	6.531	143.160	-20.300	-50.347	-50.280	-6.367
Household Attributes	Income (divided by 10000)	-0.003(-0.3)	-1.435	0.021(-0.7)	2.762		
	Kids					0.167(-2.1)	1.156
Individual attributes	Age (divided by 10)	-0.024(-2.4)	-4.008	0.083(-2.8)	3.943	0.146(-1.8)	3.348
	Student	0.036(3.7)	1.618	-0.081(2.7)	-1.014	-0.148(1.9)	-0.939
	African-American	-0.051(-4.9)	-1.728			0.523(-6.4)	1.810
	Hispanic	-0.105(-9.9)	-2.981	0.257(-8.2)	2.492	0.626(-7.6)	2.559
	Caucasian	-0.051(-4.9)	-1.728	0.129(-4.2)	1.664	0.309(-3.8)	1.477
Employment attributes	Full-time	-0.080(-7.7)	-3.751	0.281(-9.0)	4.028	0.472(-5.8)	3.294
	Flex Work	0.059(6.1)	4.234	-0.184(6.3)	-3.571	-0.372(4.8)	-3.664
	Retail	0.038(3.8)	1.725	-0.163(5.6)	-2.314	-0.195(3.0)	-1.224
	Industry	-0.031(-3.1)	-1.461			0.260(-3.2)	1.904
	Education	-0.031(-3.1)	-1.461	0.161(-5.2)	2.441		
	Office	0.022(2.3)	1.078			-0.225(2.9)	-1.64
	Government	-0.050(-4.9)	-2.107	0.161(-5.2)	2.441	0.260(-3.2)	1.904
Trip attributes	Drive alone	-0.042(-4.1)	-2.109	0.236(-7.6)	3.243	0.286(-3.6)	1.144
	Shared Ride					0.217(-2.7)	0.740
	External	0.062(6.4)	0.927	-0.23(8.0)	-0.968		
	Cost	-0.017(-1.7)	-8.649	0.064(-2.1)	9.862	0.099(-1.3)	6.338
Season	Summer	0.037(3.7)	0.97829	-0.164(5.6)	-1.223		
	τ (inverse of variance)	12.67	28.973				
	σ_v (SD of v_i)					1.285	8.169
Number of observations		1717		1717		1717	
DIC (Deviance Information Criterion)		21849.1		22559.6		19976.1	

Note: ME = Elasticity around Mean (shown parenthetically)

Table 4. Final Model Results for Work-to-Home Trips

Attribute	Variable	Lognormal		Weibull		Weibull with UH	
		Beta	t-stat	Beta	t-stat	beta	t-stat
	Alpha (α)	-	-	6.1480	17.831	6.2240	18.030
	Constant (b0)	6.7680	96.493	-43.0700	-17.580	-43.6500	-17.948
Household Attributes	Income (divided by 10000)	0.006(0.6)	1.816				
	Kids	-0.018(-1.8)	-0.656	0.073(-1.2)	0.960	0.075(-1.2)	0.986
Individual attributes	Age (divided by 10)	-0.008(-0.8)	-0.931	0.117(-1.9)	5.093	0.118(-1.9)	5.275
	Student	0.039(4.0)	1.232	-0.098(1.6)	-1.118	-0.089(1.4)	-1.013
	African-American			-0.200(3.3)	-1.253	-0.207(3.4)	-1.339
	Hispanic	0.088(9.3)	2.159	-0.122(2.0)	-1.022	-0.155(2.5)	-1.151
	Caucasian	0.059(6.1)	1.915	-0.122(2.0)	-1.022	-0.126(2.1)	-1.121
Employment attributes	Full-time	0.018(1.9)	0.681	0.085(-1.4)	1.101	0.092(-1.5)	1.254
	Flex Work	-0.013(-1.3)	-0.654	0.05(-0.8)	0.912	0.05(-0.8)	0.915
	Retail			-0.143(2.4)	-1.906	-0.144(2.4)	-1.949
	Industry	-0.044(-4.3)	-1.377	0.178(-2.9)	2.609	0.154(-2.5)	1.771
	Office	-0.024(-2.3)	-1.066				
	Government	-0.024(-2.3)	-1.066	0.178(-2.9)	2.609	0.206(-3.3)	2.387
Trip attributes	Drive alone	0.03(3.1)	0.614	-0.371(6.2)	-2.786	-0.339(5.6)	-2.598
	Shared Ride	0.042(4.3)	0.786	-0.371(6.2)	-2.786	-0.339(5.6)	-2.598
	External	0.078(8.1)	0.787	-0.663(11.4)	-2.389	-0.669(11.4)	-2.420
	Cost	0.003(0.3)	1.170	0.031(-0.5)	3.748	0.032(-0.5)	3.885
Day of week	Friday	-0.056(-5.5)	-2.169	0.122(-2.0)	1.705	0.123(-2.0)	1.694
	τ (inverse of variance)	7.4550	26.894	-	-	-	-
	σ_v (SD of v_i)	-	-	-	-	0.0506	1.660
Number of observations		1472		1472		1472	
DIC(Deviance Information Criterion)		21439.8		19667.6		19682.3	

Note: ME = Elasticity around Mean (shown parenthetically)

Table 5. Final Model Results for NHB Trips

| Variable	Attribute	Lognormal		Weibull		Weibull with UH	
		Beta	t-stat	Beta	t-stat	beta	t-stat
	Alpha (α)			4.441	55.932	4.403	56.096
	Constant (b0)	6.763	244.948	-30.7	-54.433	-30.46	-53.495
Household Attributes	Income (divided by 10000)			0.010(-0.2)	2.428	0.010(-0.2)	2.464
	Kids	-0.055(-5.3)	-4.121	0.245(-5.4)	6.106	0.241(-5.3)	5.874
Individual Attributes	Age (divided by 10)	-0.181(-16.6)	-5.106	0.106(-2.4)	10.993	0.105(-2.4)	10.78
	Male	-0.019(-1.9)	-2.01	0.077(-1.7)	2.636	0.077(-1.7)	2.632
	Hispanic	-0.065(-6.3)	-2.651			0.055(-1.2)	0.97
	Caucasian	-0.027(-2.7)	-1.471	0.0308(-0.7)	0.774	0.055(-1.2)	0.97
Employment attributes	Full-time	0.072(7.4)	5.332	-0.287(6.7)	-8.049	-0.288(6.8)	-7.919
	Part-time	0.052(5.4)	2.691	-0.100(2.3)	-1.784	-0.103(2.4)	-1.782
	Retired	0.031(3.2)	1.228				
Trip attributes	Work trip	-0.096(-9.1)	-8.696	0.527(-11.2)	14.255	0.522(-11.2)	14.021
	Shared ride	0.032(3.3)	3.122	-0.172(4)	-5.247	-0.175(4)	-5.313
	Cost			-0.012(0.3)	-1.973	-0.012(0.3)	-1.963
Day of week	Friday	0.029(3)	2.418	-0.134(3.1)	-3.554	-0.133(3.1)	-3.479
Season	Summer	-0.058(-5.7)	-1.954				
	Spring	0.021(2.1)	1.268	-0.085(1.9)	-1.878	-0.086(2)	-1.96
	τ (inverse of variance	9.3700	48.751				
	σ_v (SD of v_i)					0.0383	1.818
Number of observations		4856		4856		4856	
DIC (Deviance Information Criterion)		67943.6		65934.9		65935.8	

Note: ME = Elasticity around Mean (shown parenthetically)

Table 6. Normal Mixture Model Results for HBNW Trips

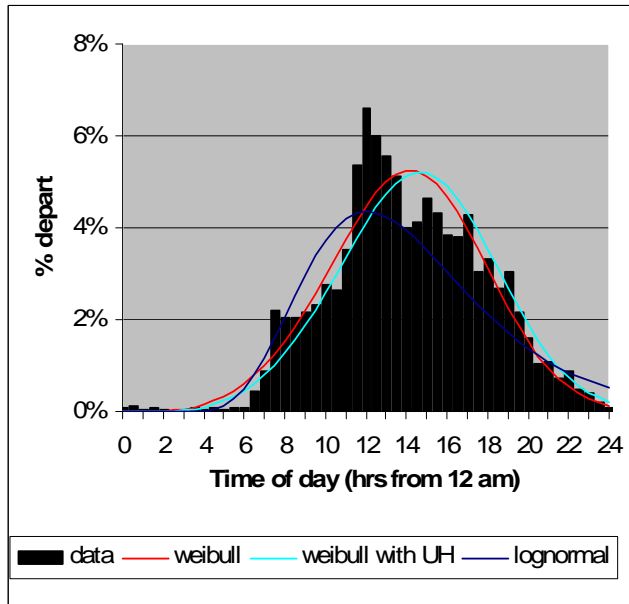| Variable	Attribute	AM Peak		PM Peak	
		Beta1	t-stat	Beta2	t-stat
	Constant (b0)	6.202	94.30	6.894	416.30
Household Attributes	Income (divided by 10000)			-0.0018	-2.45
	Kids			-0.0347	-5.00
Individual attributes	Age (divided by 100)	0.1909	2.54	-0.091	-3.85
	Student			0.01357	1.59
	Hispanic	-0.07154	-1.82	-0.0096	-0.91
	Caucasian			-0.0106	-1.23
Employment attributes	Full-time	0.07171	2.30	0.08576	8.71
	Part time	0.1297	3.78	0.05601	5.12
	Retired	0.05743	1.44	-0.0673	-3.40
	Retail	-0.08453	-2.14		
	Industry	0.05343	1.20		
	Education	0.06427	1.11	0.02036	1.44
	Office	-0.05817	-1.63		
	Government	0.05204	1.19	0.01302	1.20
Trip attributes	Drive alone	0.05448	1.69	0.08308	8.42
	Shared ride	0.05637	1.91	0.1113	13.10
	Cost	0.006153	1.49	-0.0047	-3.68
Season	Spring	0.04648	1.68		
	Summer			0.03339	2.61
P (probability)		0.3018	25.49	0.6982	58.97
τ (inverse of variance)		5.507	29.08	265.5	9.40
Number of observations		7325			

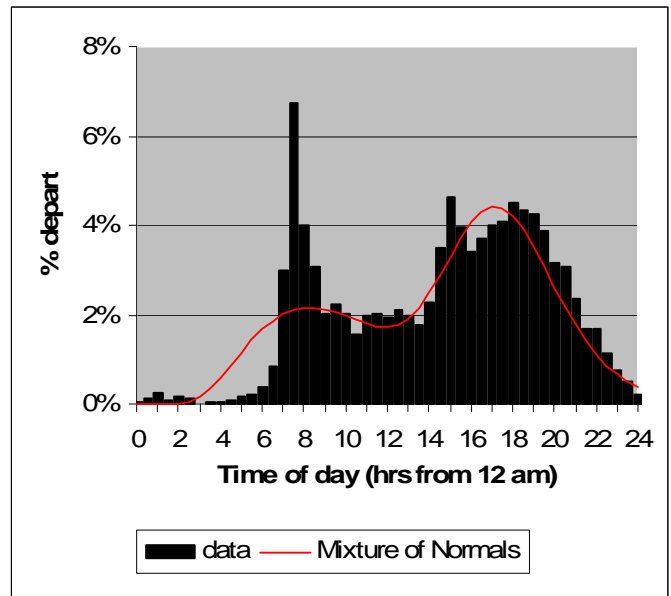Figure 1. Comparison of Model Predictions for Different Trip Purposes
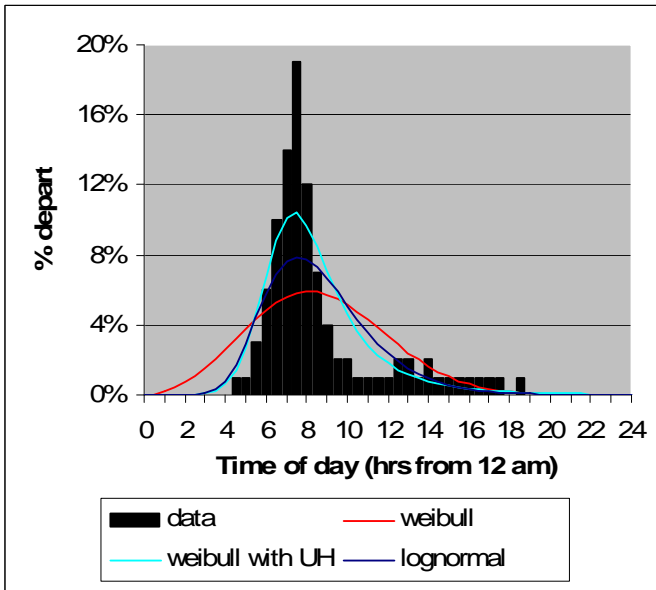
a. Home-to-Work Trips.
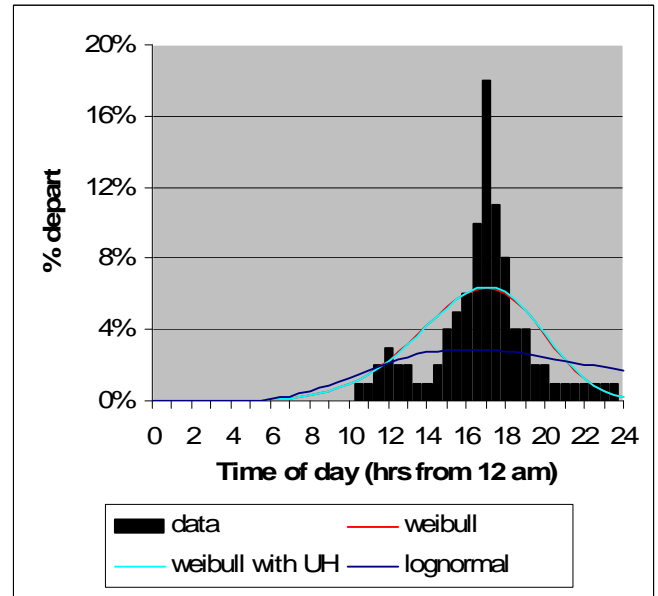
b. Work-to-Home Trips

c. NHB Trips

d. HBNW Trips